



**A University of Sussex PhD thesis**

Available online via Sussex Research Online:

<http://sro.sussex.ac.uk/>

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details

**Evaluating Development Projects: Exploring a Synthesis  
Model of the Logical Framework Approach  
and Outcome Mapping**

**Ting Yang**

**Thesis submitted for the degree of Doctor of Philosophy**

**Institute of Development Studies**

**University of Sussex**

**July 2018**

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signature: Ting Yang

University of Sussex

Ting Yang, PhD in Development Studies

Evaluating Development Projects: Exploring a Synthesis Model of the Logical  
Framework Approach and Outcome Mapping

**Summary**

Under the current results-driven development agenda, sound evaluation, and a corresponding evaluation toolkit, need to be in place to examine whether and to what extent development interventions have achieved their targeted objectives and results, and to generate lessons for further development learning and improvement. My review of the literature shows that innovative and appropriate evaluation approaches are needed to address key challenges in evaluation such as the tension between learning and accountability objectives, the need to unpack the mechanisms linking outputs and outcomes or goal, and to add an actor perspective.

Irrespective of project type, the Logical Framework Approach (LFA) is often a standard requirement of major official donor agencies on projects they fund, so as to fulfil bureaucratic imperatives. However, it is often considered inadequate in addressing key challenges in development evaluation. Given the dominant status of the LFA with such strong support from donors, it is helpful to seek a ‘middle way’: a combination of the LFA with other approaches in order to address some of its inadequacies, while satisfying donor agencies’ requirements. A synthesis of the LFA and Outcome Mapping (OM) is one such option.

This thesis explores the practical value and usefulness of a synthesis model empirically. Applying the model in two case study aid projects, I found that it serves well as a theory-based evaluation tool with a double-stranded (actor strand and results chain) theory of change. The model helps reconcile learning and accountability and add explanatory power and an explicit actor perspective. It also helps establish causation and enable attribution claims at various results levels with its different elements. The model has some limitations but my results suggest it can be usefully adopted. The choice of its application depends on project evaluation context and purpose in specific cases.

## **Acknowledgments**

During this PhD journey, I have enjoyed the support of many people in different countries and in various ways, and I would like to take this opportunity to express my sincere gratitude to them. First, I would like to thank my supervisor Dr Rosie McGee, especially for her very timely and meticulous feedback, motivation and encouragement throughout the thesis writing stage. I would also like to thank my other supervisor Dr Edoardo Masset for his kind support and guidance during earlier stages of this study, even when he was on leave of absence. My thanks also go to Institute of Development Studies (IDS) fellows Dr Martin Greeley for his help especially during my fieldwork and Ben Ramalingam for sharing his insights into my research topic.

I am most grateful to all those who offered their help in various ways during my fieldwork. I would like to give particular thanks to Dr Patrick Schroeder, Professor Zhang Mingshun and his team at Beijing University of Civil Engineering and Architecture, officers of the European Union (EU) Delegation, and respondents from all involved organizations in both Beijing and Tianjin for their kind support and cooperation during my research in China. My sincere thanks also go to Zahirul Haque, Dr Sharif Ahmed Chowdhury, Golam Mawla and his team, colleagues and respondents at Palli Karma-Sahayak Foundation (PKSF) and its partner organizations, Dr Reaz Islam, and Rakib Islam Sium for their kind support, efficient coordination and friendship during my research in Bangladesh.

I would like to give thanks to Steff Deprez, Simon Hearn, Ricardo Wilson-Grau and Kaia Ambrose for their kind attention and patience in answering my questions related

to Outcome Mapping, providing suggestions, and sharing information on the practical use of the method.

I sincerely appreciate the understanding and encouragement of my friends along this PhD journey. Last but not the least, I would like to give my utmost gratitude to my parents and siblings for their unconditional love, support and understanding, and I dedicate this thesis to them.

## Table of Contents

<b>Chapter 1 Introduction.....</b>	<b>1</b>
1.1 Background and the international development context .....	1
1.2 The broad evaluation context under the current development agenda .....	2
1.3 Research purpose and overall research design .....	5
1.4 Contribution of the thesis.....	7
1.5 Thesis structure .....	9
<b>Chapter 2 Literature review and conceptual framework .....</b>	<b>12</b>
2.1 Objectives of evaluation in the development field .....	12
2.2 Methodological and epistemological evolution of evaluation.....	15
2.2.1 Positivist and constructivist approaches to evaluation.....	16
2.2.2 Critical realism in evaluation .....	19
2.3 Challenges and gaps in development evaluation.....	22
2.3.1 Challenges in reconciling learning and accountability .....	23
2.3.2 Challenges in addressing attribution.....	25
2.4 Overview of relevant concepts .....	28
2.4.1 Learning in evaluation .....	28
2.4.2 Accountability in evaluation .....	33
2.4.3 Learning and accountability in evaluation .....	36
2.4.4 Actor-centred perspective .....	38
2.4.5 Explanatory power – unpacking the mechanisms linking outputs and outcomes or goal.....	41
2.5 Conceptual framework.....	44
<b>Chapter 3 Critical review of existing approaches in the evaluation toolkit and exploration of potential alternative .....</b>	<b>48</b>
3.1 Overview of selected approaches in the evaluation toolkit .....	48
3.2 Critical review of the Logical Framework Approach and Outcome Mapping in evaluation.....	57
3.2.1 Critical review of the Logical Framework Approach in evaluation.....	57
3.2.1.1 Overview of the LFA.....	57
3.2.1.2 Accountability and learning in evaluation with the LFA.....	60
3.2.1.3 Actor consideration with the LFA .....	62
3.2.1.4 Explanatory power of the LFA – unpacking mechanisms linking outputs and outcomes or goal in evaluation .....	63
3.2.1.5 LFA in practice – observations from selected examples .....	64
3.2.2 Critical review of Outcome Mapping in evaluation.....	68
3.2.2.1 Overview of OM.....	68
3.2.2.2 Actor consideration with OM .....	70
3.2.2.3 Accountability and learning in evaluation with OM .....	71
3.2.2.4 Explanatory power of OM – unpacking mechanisms linking outputs and outcomes or goal in evaluation .....	73
3.2.2.5 OM in practice.....	74
3.2.3 OM and the LFA .....	77
3.3 The synthesis model of the Logical Framework Approach and Outcome Mapping: a potential evaluation alternative? .....	79
<b>Chapter 4 Research Design and Methodology .....</b>	<b>85</b>
4.1 Research objective and questions .....	85
4.2 Research design and strategy.....	86
4.2.1 A case study approach (rationale and selected projects) .....	88
4.2.1.1 Rationale for case study selection .....	88
4.2.1.2 Case study projects: background and context .....	90
4.2.2 Evaluation design with the synthesis model .....	95
4.2.3 Data collection .....	99

4.2.4 Data analysis .....	103
4.2.5 Case study analysis protocol and justification of analysis strategy .....	105
<b>4.3 Reflections on the research process.....</b>	<b>107</b>
4.3.1 Reflection on limitations and challenges .....	108
4.3.1.1 Methodological reflection .....	108
4.3.1.2 Reflection on other aspects of the study .....	113
4.3.2 Reflections on ethics in the research process.....	113
<b>Chapter 5 Application of the Synthesis Model – Case 1.....</b>	<b>117</b>
<b>5.1 Case 1 .....</b>	<b>117</b>
5.1.1 Key features of Case 1 .....	117
5.1.2 Application of the synthesis model and formulation of a holistic project framework .....	119
5.1.3 Actor analysis.....	126
5.1.4 Results chain analysis .....	131
5.1.5 Evaluation mechanism analysis .....	133
<b>5.2 Uncovering the distinctive value in applying the synthesis model .....</b>	<b>138</b>
5.2.1 An extended discussion of the actor component in Case 1 .....	139
5.2.2 Linking and comparing the synthesis model with the LFA and the existing evaluation framework .....	140
5.2.3 Analysis of influencing contextual factors in Case 1 .....	142
<b>Chapter 6 Application of the Synthesis Model – Case 2.....</b>	<b>145</b>
<b>6.1 Case 2 .....</b>	<b>145</b>
6.1.1 Key features of Case 2 .....	145
6.1.2 Application of the synthesis model and formulation of a holistic project framework .....	146
6.1.3 Actor analysis.....	152
6.1.4 Results chain analysis .....	154
6.1.5 Evaluation mechanism analysis .....	155
<b>6.2 Uncovering the distinctive value in applying the synthesis model .....</b>	<b>161</b>
6.2.1 An extended discussion of the actor component in Case 2 .....	162
6.2.2 Linking and comparing the synthesis model with the LFA and the existing evaluation framework .....	164
6.2.3 Analysis of influencing contextual factors in Case 2.....	166
<b>Chapter 7 Conclusions.....</b>	<b>169</b>
<b>7.1 The synthesis model – a theory-based, mixed-methods evaluation framework...</b>	<b>169</b>
<b>7.2 Responses of the synthesis model to identified evaluation challenges.....</b>	<b>172</b>
7.2.1 Reconciling the tension between learning and accountability.....	172
7.2.2 An actor perspective in the synthesis model .....	176
7.2.3 Explanatory power of the synthesis model – unpacking mechanisms linking outputs and outcomes or goal in evaluation.....	177
<b>7.3 Additional value added and limitations arising in applying the synthesis model in practice.....</b>	<b>179</b>
7.3.1 Value added by the synthesis model.....	179
7.3.2 Limitations and challenges arising in the practical application of the synthesis model.....	182
<b>7.4 Reflections and implications .....</b>	<b>184</b>
7.4.1 Reflection and Lessons from the application of the synthesis model .....	184
7.4.2 Appropriate circumstances for using the synthesis model .....	188
7.4.3 Further areas for research on the synthesis model .....	189
<b>Bibliography .....</b>	<b>191</b>
<b>Appendices.....</b>	<b>205</b>
Appendix 1: Glossary of Outcome Mapping terms (definition of elements and steps). 205	



Appendix 2: Logic model of the Logical Framework Approach .....	206
Appendix 3: Logic model of Outcome Mapping .....	207
Appendix 4: Framework used to collect progress-marker-related information .....	208
Appendix 5: Interview guide – boundary partners: local consumer associations.....	209
(Case 1) .....	209
Appendix 6: Interview guide – boundary partners: retailers (Case 1) .....	211
Appendix 7: Interview guide – boundary partners: suppliers (Case 1).....	213
Appendix 8: Interview guide – boundary partners: local authorities (Case 1) .....	215
Appendix 9: Interview guide – project team staff (Case 1) .....	217
Appendix 10: Interview guide – donor representatives (Case 1).....	217
Appendix 11: Interview guide – boundary partners: partner organizations.....	218
(Case 2) .....	218
Appendix 12: Interview guide – project team staff (Case 2) .....	220
Appendix 13: Interview guide – staff from the donor agency (Case 2).....	221

## List of Tables

<i>Table 2.1: Primary dimensions of contrasting methodological paradigms.....</i>	<i>18</i>
<i>Table 3.1: A typology of evaluation models.....</i>	<i>49</i>
<i>Table 3.2: A critical review and analysis of eight selected approaches .....</i>	<i>51</i>
<i>Table 3.3: The Logical Framework Matrix .....</i>	<i>57</i>
<i>Table 3.4: Hypothesis and forms of evaluation in LFA-based project and evaluation design</i>	<i>60</i>
<i>Table 3.5: Case studies of evaluation with the LFA.....</i>	<i>65</i>
<i>Table 3.6: Case studies of evaluation with OM.....</i>	<i>74</i>
<i>Table 3.7: LFA and OM in evaluation.....</i>	<i>77</i>
<i>Table 4.1: Reformulated project matrix of the synthesis model.....</i>	<i>98</i>
<i>Table 4.2: A brief summary of interviewees in the two case studies.....</i>	<i>100</i>
<i>Table 5.1: Key features of Case 1 .....</i>	<i>118</i>
<i>Table 5.2: Case 1 reformulation framework.....</i>	<i>120</i>
<i>Table 5.3: Boundary partners – representative retailers’ achievement of progress markers</i>	<i>136</i>
<i>Table 5.4: Boundary partners – local consumer associations’ achievement of progress markers .....</i>	<i>137</i>
<i>Table 6.1: Key features of Case 2 .....</i>	<i>145</i>
<i>Table 6.2: Case 2 reformulation framework.....</i>	<i>147</i>
<i>Table 6.3: Self-sufficiency ratios of partner organization branches by year.....</i>	<i>165</i>
<i>Table 7.1: Example of the contribution of OM elements to the learning pathway .....</i>	<i>174</i>

## List of Figures

<i>Figure 2.1 Conceptual framework.....</i>	<i>46</i>
<i>Figure 3.1: Three stages of OM.....</i>	<i>69</i>
<i>Figure 3.2: Sphere of project influence .....</i>	<i>73</i>
<i>Figure 3.4: Synthesis model of LFA and OM.....</i>	<i>80</i>
<i>Figure 4.1: Overall research design and strategy (a case study approach).....</i>	<i>88</i>
<i>Figure 4.2: Seven components of PRIME.....</i>	<i>94</i>
<i>Figure 5.1: Project sphere of influence .....</i>	<i>127</i>
<i>Figure 5.2: Actor strand of the theory of change in Case 1 .....</i>	<i>128</i>
<i>Figure 5.3: Results strand of the theory of change in Case 1 .....</i>	<i>131</i>
<i>Figure 5.4: Analysis of influencing factors in Case 1 .....</i>	<i>143</i>
<i>Figure 6.1: Actor strand of the theory of change in Case 2 .....</i>	<i>153</i>
<i>Figure 6.2: Results strand of the theory of change in Case 2 .....</i>	<i>154</i>
<i>Figure 6.3: Changes in income and expenditure of ultra-poor.....</i>	<i>157</i>
<i>households by year .....</i>	<i>157</i>
<i>Figure 6.4: Project households at different levels of food.....</i>	<i>157</i>
<i>security during Monga by year.....</i>	<i>157</i>
<i>Figure 6.5: Partner organization ATOM score .....</i>	<i>158</i>
<i>Figure 6.6: Partner organization ESSAS.....</i>	<i>159</i>
<i>Figure 6.7: Frequently mentioned key words in partner organization interviews .....</i>	<i>161</i>
<i>Figure 6.8: Analysis of influencing factors in Case 2 .....</i>	<i>166</i>
<i>Figure 7.1: Double-stranded theory of change under the synthesis model.....</i>	<i>170</i>

## List of Abbreviations

ADB	Asian Development Bank
AGRIDEA	Swiss Association for the Development of Agriculture and Rural Areas
AusAID	Australian Agency for International Development
BUCEA	Beijing University of Civil Engineering and Architecture
CA	Contribution Analysis
CBA	Cost-Benefit Analysis
CEDIL	Centre of Excellence for Development Impact and Learning
CIDA	Canadian International Development Agency
DFID	Department for International Development
EC	European Commission
ECODE	Development Cooperation Study (abbreviation in Spanish)
ETH	Swiss Federal Institute of Technology
EU	European Union
FASID	Foundation for Advanced Studies on International Development
GTZ	German Organization for Technical Cooperation
ICAI	Independent Commission for Aid Impact
IDRC	International Development Research Centre
INTRAC	International NGO Training and Research Centre
JICA	Japan International Cooperation Agency
LF	Logical Framework
LFA	Logical Framework Approach
MDG	Millennium Development Goal
MoU	Memorandum of Understanding
NGO	Non-Governmental Organization
NORAD	Norwegian Agency for Development Cooperation
NPM	New Public Management
ODI	Overseas Development Institute
OECD	Organization for Economic Cooperation and Development
OECD-DAC	Development Assistance Committee of the Organization for Economic Cooperation and Development
OM	Outcome Mapping

PEA	Political Economy Analysis
PME	Planning, Monitoring and Evaluation
RBM	Results-Based Management
RBM	Results-Based Monitoring
RCT	Randomized Controlled Trial
ROM	Results-Oriented Monitoring
SC	Sustainable Consumption
SDC	Swiss Agency for Development and Cooperation
SDG	Sustainable Development Goal
Sida	Swedish International Development Cooperation Agency
SME	Small and Medium-sized Enterprise
ToC	Theory of Change
UK	United Kingdom
UN	United Nations
USAID	United States Agency for International Development
VA	Voluntary Agreement
VfM	Value for Money

## **Chapter 1 Introduction**

### **1.1 Background and the international development context**

With the pace of globalization, the changing socio-economic environment, and the influence of New Public Management (NPM) theories that dominate modern public sector management (Earle, 2003; Kilby, 2004), new trends and demands have been raised onto the international development agenda, such as the results-based approach to development programme management. Specifically, during the 1990s, public sector reforms took place in many countries – especially members of the Organization for Economic Co-operation and Development (OECD) – to respond to economic, political and social pressures (OECD, 2000). These reforms shifted the focus of government efforts from tracking resources and activities to improving performance and ensuring the achievement of desired results – performance management or results-based management (RBM) (ibid.; OECD, 2014).

In such a context, the Millennium Development Goals (MDGs) were initiated and agreed by government leaders in 2000 as ‘one of the first attempts to focus development cooperation on quantifiable results rather than monitoring inputs’ (OECD, 2014, p.15). Further strengthened with the endorsement of the Paris Declaration on Aid Effectiveness in 2005 and the Sustainable Development Goals (SDGs) in 2015, ‘focusing on results’ has become a cornerstone of international development cooperation (OECD, 2014; Roduner et al., 2008; United Nations, 2015).

Under the results-focused development agenda set through these global endeavours, there has been increasing demand to demonstrate achievement of development results with the emphasis on performance measurement (OECD, 2000). The latter concerns performance information gathering and the technical aspects of articulating and

clarifying objectives, selecting and developing indicators, and collecting and analysing data on results (ibid.). A logic model such as the logical framework is often included in the performance measurement system at operational level to show linkages between objectives and results (Armytage, 2011).

Meanwhile, the nature and role of aid in development, and the relationship between donors and recipients have been undergoing significant changes (Stern et al., 2012). Thus, the development context is becoming increasingly complex with the influences of diversified actors and other drivers of development results (ibid.; Conlin and Stirrat, 2008).

## **1.2 The broad evaluation context under the current development agenda**

In the development context outlined above, sound evaluation and a corresponding evaluation toolkit are needed to examine whether and to what extent development efforts have achieved targeted objectives, and to generate lessons for further learning and improvement. Under the current development agenda, evaluation has become considerably more important and has been repositioned centrally in the development field as a means of promoting development effectiveness and achieving development goals (Armytage, 2011). It is believed that ‘in the future, evaluation will be an agent of change for the world’s 193 nations committed to achieving the 2030 Sustainable Development Agenda’ (Segone, 2016, p.4). Accordingly, evaluation will face greater complexity with the interconnected and synergic goals and targets of the SDGs and their inclusive focus on leaving no one behind (ibid.).

Apart from the need for rigorous evaluation in measuring the achievement of macro-

level development goals agreed collectively by the international community, there is also the longstanding need to assess the effectiveness of interventions at the project level. And there is a newer demand to assess effectiveness of initiatives designed as complex, multi-year, often multi-site and multi-donor programmes, in which the usual challenges of project-level evaluation are compounded by the need to make sense of changes happening at the overall programme level in relation to what goes on at the micro level in its individual components. As the focus of development evaluation under the results agenda shifts from measuring outputs to outcomes and impacts, evaluation needs to adapt to reflect these changes (Bamberger et al., 2009).

With the increasingly complex development context, and evolving nature and changing landscape of development aid and interventions, Stern et al. (2012, p.11) summarize some of the emerging demands: 1) Evaluation design should be appropriate to programme and project features – often complex, implemented indirectly through agents, involving multiple partners, and only comprising a small part of the wider development portfolio. 2) Evaluation needs to answer multiple questions ranging from the basic – ‘did it work?’, and explanatory – ‘how did it work?’ to more elaborate judgement – ‘for whom do interventions make a difference?’. 3) Evaluation should be able to address multi-dimensional and overlapping development interventions with (often) long-term ambitions. 4) Evaluation should be adapted to devolved partnerships and trust-based relationships with aid recipients, which, at the same time, limits what donors can evaluate on their own.

While the consensus around these emerging demands is strong, however, the question of how to meet them is less resolved. Some widely recognised evaluation challenges



give rise to differences in view on this, including the difficulty involved in achieving both key objectives of learning and accountability (Cracknell, 2000), and the problem of attribution that is caused by the complexity of the development process and multiple influencing factors at play (Conlin and Stirrat, 2008; White, 2010). To resolve these challenges, appropriate and innovative evaluation approaches and methods with certain explanatory power and an actor perspective are needed. However, as White (2005, p.iii) observes, ‘there has been inadequate investment in methodology, often resulting in low quality evaluation outputs’ and ‘attention to sound methodology matters’. This observation remains valid. For example, the Methods Lab initiative (ODI, 2018) and the establishment of the Centre of Excellence for Development Impact and Learning (CEDIL) (CEDIL, 2018) are examples showing attention to and investment in evaluation methodologies. The ever-changing development context and the complex and dynamic nature of interventions involving diverse actors and factors, call for the constant evolution of evaluation techniques and tools in order to meet changing needs and address key challenges. Continuously enriching the methodology toolkit and database requires the continuous devotion of research and effort to exploring such possibilities and experimenting with innovative approaches to test their feasibility and applicability empirically.

The Logical Framework Approach (LFA), the most widely adopted project planning, monitoring and evaluation tool, is often considered inadequate to address key challenges in development evaluation. Yet, in practice, it is imposed by major official donor agencies on most interventions to fulfil bureaucratic imperatives irrespective of its suitability for project type. Therefore, given the dominant status of the LFA, it seems expedient to explore a middle way, possibly synthesizing it with other approaches to address some of its inadequacies while satisfying donor agency

requirements.

A theoretically proposed synthesis model of the LFA and Outcome Mapping (OM) appears to be one such option (Roduner et al., 2008). This model seeks to combine the strengths and advantages of the two approaches through integration of some of their key elements. Methodologically, the model is assumed to have particular potential value in addressing contemporary evaluation challenges and reflecting the complexity of the development process. Epistemologically, it is considered to represent critical realist principles in evaluation and strive for a balance between the inadequacies of the positivist oriented LFA and constructivist oriented OM (Kontinen, 2010; Patton, 2008; White and Phillips, 2012). Nevertheless, this synthesis model was proposed theoretically. It needs to be tested in practice, experimented with, and examined for its applicability and practical value in development evaluation in reality. I have tried to ascertain how far the model has been applied in evaluation in practice and identify written-up cases in this respect but have not been able to identify any, even through contacting the original author of the model and extensive literature review.

### **1.3 Research purpose and overall research design**

Grounded in the current development agenda and broader evaluation context, this research project aims to further explore and pursue adaptive and feasible alternatives for evaluating development interventions. Specifically, I put into practice and tested the aforementioned theoretical synthesis model of the LFA and OM in development evaluation. The major objectives were to examine empirically the applicability and value of the model as an evaluation tool and to draw useful lessons and implications for future reference and research. The main research question and sub-questions are as follows:

What evaluative insights are afforded by integrating aspects of the LFA and OM to construct a synthesis model?

- 1) How does the synthesis model, as an alternative evaluation approach, respond to identified evaluation challenges and demands – in particular how well does it reconcile the tension between learning and accountability, offer explanatory power about the links between outputs and outcomes or goal and add an actor perspective?
- 2) What value is added and what limitations arise in adopting the synthesis model in practice?
- 3) What are the lessons and implications for further research and evaluation practice?

To serve the research purpose and answer research questions, I adopted a case study strategy and operationalized my research activities by testing the synthesis model with two aid projects: a European Union (EU) funded intervention on sustainable consumption in China; and an aid project in Bangladesh on poverty reduction funded by the United Kingdom (UK) Department for International Development (DFID). These two case studies were conducted with the aim of unpacking the evaluation function of the model through making explicit its underlying theories of change and evaluation mechanisms, and exemplifying its usage. During my fieldwork research, I acted as an independent evaluator and conducted external, non-official evaluations of the two projects using the synthesis model. Given the limited scope and capacity of a PhD research project, the synthesis model was applied directly as a post-implementation evaluation tool rather than earlier in the implementation of the case study projects.

## **1.4 Contribution of the thesis**

This thesis contributes to the evaluation literature and practice, especially in the development field, in terms of several aspects. First, the thesis seeks to provide an in-depth review of key contemporary evaluation challenges, issues and related concepts. To this end, it articulates a conceptual framework based on these concepts which reflects the key areas of concern in evaluation. This conceptual framework provides an analytical anchor with key parameters – reconciling the tension between learning and accountability, offering explanatory power through unpacking the mechanisms linking outputs and outcomes or goal, and adding an actor perspective – for assessing the evaluation approaches explored in this study and generating valuable insights on existing methodological gaps and potential solutions.

Insights gathered through a secondary review of the merits and demerits of a range of evaluation approaches using the aforementioned framework help enrich the evaluation literature, and methodological gaps and potential solutions identified may help provide some useful references for evaluators in practice.

Second, this thesis contributes to evaluation literature and practice by exploring a potential alternative evaluation approach which seeks to synthesize various key dimensions and aspects of evaluation into one framework. It does so by providing empirical cases of the application of the theoretical synthesis model in evaluating two development interventions supported by major official donor agencies. This study 1) helps enlarge the geographical coverage of OM application to China where it is scarcely used; 2) applies a critical research perspective to a case of evaluation practice entailing the application of a relatively untried approach; and 3) helps build the case for the critical realist evaluation paradigm insofar as the applied model, which is

critical realist in orientation, and demonstrates its usefulness. In addition, this study explores and exemplifies the possibility of utilizing and analysing evaluation data to illustrate the usefulness of a theoretically proposed model, through unpacking the evaluation mechanism and process embedded in the model from a combined ‘actor’ and ‘results’ perspective. In all of these respects, the study has applications in the development evaluation field beyond the scope of this PhD project.

While these two case studies do not lend themselves to large-scale generalization, they can support ‘under these conditions’ generalization (Stern et al., 2012). Indeed, they represent two different case-specific examples of the application of the synthesis model, which provides lessons and reference points for other projects implemented under similar conditions, or with similar features, so that a small cluster of cases can be gradually built up for future research. Experiences gathered and lessons learnt from applying the synthesis model in this research project through reflections on various key issues in the research process can also serve as useful reference for future application of the model.

Third, some observations generated from this thesis can help inform some areas for practical improvement or future research. For example, the empirical research of this study revealed the need for transformation in mindset from ‘demand’ – recipient or beneficiary – to ‘supply’ side – service provider/project implementer – especially among frontline project staff. This may have implications for management decision-making in relation to awareness-raising and capacity-building support to relevant project members. Finally, the challenges and achievements of this study will also benefit future researchers interested in further exploration of this field. For example, this thesis points to the necessity of an actor-centred research on donor receptivity and

perceptions of the synthesis model given strong influence of donor recognition of specific evaluation approaches.

## **1.5 Thesis structure**

This thesis has seven chapters. Chapter 2 seeks to set the scene by outlining the broader evaluation context for this research project from the perspective of epistemological and methodological debate and development. It identifies and reviews key contemporary challenges and demands in evaluation and the need for pursuing innovative approaches to respond to these challenges; and articulates a conceptual framework on the basis of some key concepts drawn from this literature review – which provides key parameters in reviewing and analysing evaluation approaches involved in this thesis. First, the chapter briefly reviews development evaluation including its definition and key goals. It then presents the methodological and epistemological evolution of evaluation history, followed by identification of key challenges and demands in relation to evaluation methods. The chapter proceeds to provide two conceptual sections (sections 2.4 and 2.5) by first reviewing some key concepts drawn from the previous literature review especially in relation to these identified evaluation challenges, and then formulating a conceptual framework with these reviewed concepts.

Chapter 3 elaborates on the position set out in the previous chapter that there are some methodological gaps in the evaluation toolkit in respect of certain tensions and dilemmas prevalent in contemporary evaluation debates and practice. The chapter provides a critical review and secondary analysis of selected approaches currently available in the evaluation toolkit with reference to the key parameters in the conceptual framework constructed in Chapter 2. Specifically, this chapter first

presents a general overview and analysis of some approaches in the evaluation toolkit. It then focuses on the LFA and OM as two key evaluation methodologies and critically reviews their role in evaluation in a detailed manner. The chapter finally introduces a theoretically proposed synthesis model of the LFA and OM as an alternative evaluation approach which potentially adds value in terms of responding to certain key evaluation challenges and demands, as featured in the conceptual framework.

Chapter 4 first outlines the general research objective and questions. It then introduces in detail the overall research methodology, including utilization of a case study strategy; project selection; evaluation design with the theoretical synthesis model; primary and secondary data collection and analysis process; and justification for the case study analysis strategy adopted (which focuses on actor analysis, results chain analysis, and evaluation mechanism analysis). The chapter concludes with a critical reflection on the whole research process including limitations and challenges, methodological aspects and ethical issues.

Chapters 5 and 6 respectively begin by presenting the process of applying the synthesis model to one of the case study projects. The two chapters then proceed to unpack the key components and evaluation function of the model in each case, guided by the case study analysis strategy set out in Chapter 4. Finally, these chapters extend the discussion to a within-case analysis to further reveal the added value of using the synthesis model, and to show some case-specific contextual conditions as references for future application of the model. Chapter 6 concludes with a summary of the observed distinctive features and added value of the synthesis model during its practical application in the two case studies.

Chapter 7 reflects at a more general level on the synthesis model and its practical application, answers the research questions set out in Chapter 4, and offers various evaluative insights afforded by constructing and applying the model. The chapter first reflects on the synthesis model as a theory-based evaluation framework accommodating a mixed methods design. It then summarizes the responses of the model to contemporary evaluation challenges and demands, as reflected in the conceptual framework, the practical value it can add, and its limitations. Finally, the chapter considers the application of the synthesis model in the two case studies, and appropriate conditions and circumstances for using the model. It concludes with lessons learnt and implications for further research and evaluation practice.



## **Chapter 2 Literature review and conceptual framework**

This chapter starts with a brief introduction to development evaluation including its definition and key goals. Section 2.2 provides a review of methodological and epistemological evolution in evaluation history. Section 2.3 proceeds to identify key evaluation challenges and gaps related to evaluation methods. Section 2.4 reviews some key concepts drawn from the literature review in previous sections especially in relation to these identified evaluation challenges and demands. Section 2.5 then articulates a conceptual framework with these reviewed concepts.

### **2.1 Objectives of evaluation in the development field**

‘Evaluation is a very young discipline – although it is a very old practice’ (Scriven, 1996, p.395); and, as Pawson and Tilley (1997, p.1) point out, ‘it has become axiomatic as we move towards the millennium that everything, but everything, needs evaluating’. Indeed, evaluation presents itself in many different fields and there are various efforts to define it from different perspectives and focuses, such as value (Fournier, 2005), utilization (Patton, 2008), and an application of social research methods (Rossi, 2004).

The following entry in the *Encyclopedia of Evaluation* is one relatively widely used generic definition:

Evaluation is an applied inquiry process for collecting and synthesizing evidence that culminates in conclusions about the state of affairs, value, merit, worth, significance, or quality of a program, product, person, policy, proposal, or plan. (Fournier, 2005, p.140)

It can thus be seen that the objects of evaluation are diverse and range from programmes, policies and people to products and so on. Indeed, the ‘early knowledge

base of evaluation borrowed heavily from existing methods and theories in nearly all social sciences' (Shadish et al., 1991, p.28), and has been evolving and becoming more specialized ever since. Since the beginning of the 1960s, evaluation gradually grew and flourished as a profession with the accumulation of knowledge and the creation of professional publications, societies and professional standards or codes of conduct (Shadish et al., 1991; Shadish and Luellen, 2005).<sup>1</sup>

From a more general perspective, evaluation is regarded by Guba and Lincoln (1989) as having gone through four 'generations' of evolution with a shift in focus from measurement, description, and judgement of worth to focus on a process of negotiation about an agreed action plan among various stakeholders. This perspective on the development of evaluation has been well received and cited in the evaluation literature. Underpinning such a model of evaluation development, there was a corresponding epistemological evolution.

From the 1980s onwards, under the influence of the New Public Management theories (Kilby, 2004), accountability became a central element of performance management, and value for money and cost-effectiveness became key concepts in evaluation. Then around the mid-1990s, an evidence-based wave of evaluation appeared with the emphasis on real empirical evidence showing what worked. This was seen as a renaissance of science and randomized experimentation (Vedung, 2010).

In the evolutionary history of evaluation, development evaluation emerged with the post–World War II reconstruction and development work as a sub-discipline based on the audit and social science traditions of the discipline (Morra Imas and Rist, 2009).

---

<sup>1</sup> This thesis is mainly limited to a brief overview of evaluation history since the 1960s when the field started to become professionalized.

Apart from sharing some of the general characteristics of evaluation, development evaluation is specifically defined by the Development Assistance Committee of the Organization for Economic Cooperation and Development (OECD-DAC) as

the systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfillment of objectives, development efficiency, effectiveness, impact and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process of both recipients and donors. (OECD-DAC, 2002, pp.21–22)

This widely adopted definition provides the standard criteria for the practice of development evaluation, that is, efficiency, effectiveness, impact and sustainability; and also identifies learning as one of the key purposes of evaluation. More specifically, as specified by OECD-DAC (1991, p.5), the purpose of evaluation in development is to ‘improve future aid policy and programmes through feedback of lessons learned; and to provide a basis for accountability, including the provision of information to the public.’ There is widespread endorsement of this view and amongst the various potential purposes of evaluation, learning and accountability are undoubtedly the two most fundamental ones (Armytage, 2011; Cracknell, 2000). Additionally, Picciotto (2002) observes a similar status of learning and accountability as the two major goals of evaluation in the practice of the World Bank evaluation office:

In a nutshell, feedback and follow-up were central to the conception of [the Office of Evaluation and Development] OED from day one. Feedback is about learning and follow-up is about accountability. They are two sides of the same coin. (Picciotto, 2002, p.1)

However, realizing learning and accountability concurrently in evaluation is challenging. While many stakeholders in development evaluation endorse the

importance of both and affirm their compatibility, pursuing both in practice is often a matter of reconciling tensions. Such tension often influences how evaluation is organized and, while desirable, it is seldom possible to 'kill the two birds with one stone' (Cracknell, 2000, p.55). Given the importance of these two objectives but the difficulty involved in achieving both of them, evaluation practitioners and researchers have been seeking solutions through the pursuit of innovative and adaptive approaches and methods. This is also the underlying rationale of the present study.

Following this brief overview of the definition and key goals of evaluation, especially related to development, the next section focuses on the review of the methodological and epistemological evolution of evaluation and identifies relevant debates and tensions in a more detailed manner to lay out the context in which this research project was undertaken.

## **2.2 Methodological and epistemological evolution of evaluation**

There is growing recognition that evaluation can contribute to the greater effectiveness of development interventions and has become an established and powerful tool to help improve the way organizations achieve results (Morra Imas and Rist, 2009; Picciotto, 2007; Thomas, 2010). It is thus envisaged that with a new and broader agenda set through the SDGs, evaluation will play a greater role and help achieve better development results (Segone, 2016).

However, this crucial role of evaluation cannot be realized if its practice is ill-equipped: availability and the use of appropriate tools and methods are key to the overall effectiveness of evaluation. As observed earlier, the field has been evolving and tools and methods in the evaluation toolkit have been changing under the

different waves of evaluation. It is thus worth reviewing the history of this methodological evolution through a review of the underlying debates and epistemological stances.

As suggested in the previous section, although there are different perspectives on the history of evaluation, they nevertheless concur in respect of broader underlying epistemological and methodological evolution and tendencies. Since the 1960s, evaluation research and evaluative thinking have been influenced by major epistemological waves in the social sciences, including positivism, constructivism and critical/scientific realism. Corresponding with these major methodological debates there has been evolution in the use of quantitative, qualitative and mixed methods approaches to evaluation (Patton, 2008; Pawson and Tilley, 1997; Shadish, et al., 1991).

### **2.2.1 Positivist and constructivist approaches to evaluation**

The epistemological debate in the field of evaluation has traditionally been centred on a dichotomy between positivism and constructivism. In a more general sense, these two stances differ fundamentally in their conceptualization of reality, the aim and means of knowledge inquiry, the relationship between the researcher and the researched, as well as other key features (Moses and Knutsen, 2007; Sumner and Tribe, 2008). Such a dichotomy essentially concerns the ‘issue of how evaluation finds truth and contributes to knowledge’ (Armytage, 2011, p.270).

Positivists favour the uncovering of universal laws – empirical generalizations – through a scientific approach with the focus on reliability and validity of data, experimental and statistically sound methods are thus preferred under positivism

(ibid.; Patton, 2008). Due to its attractive ‘offer to provide a much-needed assurance of certainty to the business of development’ (Armytage, 2011, p.271), the positivist-oriented approach epitomized in experimental design has been promoted as a gold standard by some development practitioners (e.g. Stern et al., 2012), especially when trying to demonstrate attribution in fulfilling the evaluation goal (Rossi et al., 2004).

However, as Patton (2002, p.50) argues, to ‘claim the mantle of objectivity in the post-modern age is to expose oneself as embarrassingly naive’. Such criticism of positivist principles and the practical difficulty in applying a positivist-oriented approach have driven the quest for an alternative paradigm in development evaluation. In contrast, constructivism recognizes diversity, complexity and contextual factors in the social world, and subsequently disproves the ‘false assurance of positivist evaluation’ (Armytage, 2011, p.272). The constructivist paradigm advocates empowerment and a pro-poor, participatory stakeholder evaluation model (Cracknell, 2000), as well as evaluation criteria that ‘substitute credibility for internal validity, transferability for external validity, dependability for reliability, and confirmability for objectivity’ (Armytage, 2011, p. 272).

Significantly, constructivists consider patterns of interest to be the product of people’s own construction, a position that implies an actor-centred approach to development and change. As Long (2001) observes, development intervention can be seen as a transforming process during which external factors enter the existing lifeworlds of those groups and individuals involved and then become internalized or localized. It thus has largely different implications for different interest groups or actors regardless of their respective roles in the intervention. Under such circumstances, an actor-centred approach could offer ‘valuable insights into these processes of social

construction and reconstruction. It also enables one to conceptualize how small-scale interactional settings or locales interlock with wider frameworks, resource fields and networks of relations, thus facilitating a re-thinking of key concepts such as ‘constraints’, ‘structure’ and ‘micro–macro’ relations’ (ibid., p.49).

As constructivism has gained ground in evaluation approaches, an actor orientation has acquired increased relevance. The question evaluators face is how to achieve it without falling into excessive relativism or subjectivism and becoming unable to generate any completely clear and unambiguous evidence of or claims to impact.

The epistemological evolution of and debate around evaluation have resulted in corresponding methodological development and trends. Accordingly, Table 2.1 exemplifies the conventional methodological distinction between qualitative and quantitative approaches with their respective underlying epistemological stances of constructivism and positivism.

**Table 2.1: Primary dimensions of contrasting methodological paradigms**

<b>Qualitative Paradigm</b>	<b>Quantitative/Experimental Paradigm</b>
Qualitative data (narratives, description, quotations)	Quantitative data (numbers, statistics)
Naturalistic inquiry (openness)	Experimental design (control)
In-depth case studies	Treatment and control groups
Inductive analysis	Deductive hypothesis testing
Subjective perspective valued	Objectivity
Close and direct observation of the programme	Distant from and independent of the programme
Holistic contextual portrayal	Independent and dependent variables
Systems perspective focused on interdependencies	Linear, sequential modelling
Dynamic, continuous view of change	Pre- and post-measurement of change
Purposeful sampling of relevant cases	Probabilistic, random sampling
Focus on uniqueness and diversity	Standardized, uniform procedures
Emergent, flexible designs	Fixed, controlled design protocols
Thematic content analysis	Statistical analysis
Value uniqueness, particularity	Replication

Extrapolation (lessons and principles)	Generalization (empirically based external validity)
Constructivism/ Phenomenology	Positivism

(Source: Patton, 2008, p.433)

As a consequence of the differences but also complementarities between qualitative and quantitative approaches, the field has seen an increase in pluralist or mixed-methods evaluation design, with its combination of qualitative and quantitative tools. Pawson and Tilley (1997, p.25) argue that there is a need for ‘breadth and then depth in programme evaluation’. This is particularly necessary in the evaluation of large and complex development interventions that seeks both learning and accountability and often has theories of change and action combining quite diverse change actors and change actions. Such evaluations usually require mixed methods and both qualitative and quantitative data.

### **2.2.2 Critical realism in evaluation**

Critical realism – sometimes used interchangeably with the terms scientific realism, transcendental realism, or relational realism – is regarded as a combination of some of the most distinctive features of the positivist and constructivist approaches (Moses and Knutsen, 2007). Or according to Danermark, et al. (2002, p.202), from a methodological viewpoint, ‘critical realism constitutes a ‘third way’ in the scientific debate between, on the one hand empiricism/objectivism, and on the other hand relativism/idealism’.

More specifically, at its ontological core, critical realism shares with positivism the recognition of the existence of an independent real world (Moses and Knutsen, 2007) but recognizes many layers and domains of reality: the empirical (experience), the actual (events), and the real (whatever exists or the realm of objects and their



structures and powers) (Danermark, et al., 2002). Similarly to constructivism, critical realism realizes that the social world is full of complexity and believes that the best way to discover truths is through a scientific approach in positivist terms such as statistical methodology (Moses and Knutsen, 2007).

Critical realism builds on the assumption that objects in reality have causal powers or generative mechanisms (Pawson and Tilley, 1997; Sayer, 2000). From a realist perspective, mechanisms refer to those ‘underlying entities, processes, or structures which operate in particular contexts to generate outcomes of interest’ (Astbury and Leeuw, 2010, p.368). As social science researchers usually work in an open system which interacts and exchanges with its outside environment (Williams and Imam, 2007), the generative mechanisms under study thus ‘operate in a complex interaction with other mechanisms, which either cooperate with or work against the mechanism in question’ (Danermark, et al., 2002, p.199). Similarly, as Sayer (2000, p.5) argues, critical realists focus on ‘necessity and contingency rather than regularity, on open rather than closed systems, on the ways in which causal processes could produce quite different results in different contexts’.

Methodologically, in contrast to positivism and constructivism, critical realism accommodates a relatively wider range of research models and rejects formulaic prescriptions for method. This implies that specific method choices should be made according to ‘the nature of the object of study and what one wants to learn about it’ (Sayer, 2000, p.19).

After its manifestation in modern philosophy and the social sciences for quite some time, critical realism entered the evaluation field with the contribution of Pawson and

Tilley (1997), who were the first to explicitly import realist principles to evaluation research and proposed a methodological model: mechanism + context = outcome. Pawson and Tilley (1997, p.66) argue that a focus on programme mechanism represents an advance from ‘asking whether a programme works to understanding what it is about a programme which makes it work’.

Mechanisms usually remain unobservable or hidden and sensitive to variations in context (Astbury and Leeuw, 2010, p.368). It is argued that evaluators must determine the substantive relationship between context and mechanism in order to understand the outcome of an initiative or intervention (ibid.; Pawson and Tilley, 1997). Programmes do not operate in a vacuum but in an open and dynamic system in which context plays an enormous role in shaping them and their effects. The causal powers or generative mechanisms of a programme in different contexts can produce quite different outcomes. For example, an intervention affects and is affected by social, political and economic factors such as external funding sources, political pressure, stakeholder expectations, the capacity of local partners, and the political and economic values of society (Pawson and Tilley, 1997; Sayer, 2000; Shadish et al., 1991, p.38).

As well as placing ‘mechanism’ at the centre of critical realist evaluation methods, Pawson and Tilley (1997) emphasize the necessity of considering the nature of what it is we are evaluating. The object of study or evaluation is ‘constituted in complex processes of human understanding and interaction...whatever the program, in whatever the circumstances, it will ‘work’ through a process of reasoning, change, influence, negotiation, battle of wills, persuasion, choice increase (or decrease), arbitration or some such like’ (ibid., p.17).

In this connection, given the complex nature of the evaluand in the development field and its surrounding context of a multitude of influencing factors and actors, any evaluation should take into consideration both the causal mechanisms and specific conditions under which a programme is implemented. Constructivist (or critical realist) evaluators hold that this expectation cannot be realized and reflected by quantitative indicators alone, and that they need to be supplemented by qualitative and/or intensive research if sufficient in-depth data is to be gathered (Sayer, 2000; Pawson and Tilley, 1997). In other words, they propose that mixed methods approaches should be considered as the best options for revealing a relatively more comprehensive picture in evaluation.

Reviewing the background to evaluation in terms of methodological and epistemological evolution helped me to better understand the phenomenon and ground my research in the broader evaluation context, as well as to identify gaps in knowledge and needs for further effort. Development evaluation, which evolved out of the general evaluation field as a sub-discipline, fits into the above-mentioned broader debate and context. The following section then proceeds to review the challenges and gaps in the field.

### **2.3 Challenges and gaps in development evaluation**

As observed in Chapter 1, the international development context has constantly changed along with the evolving development agendas of the MDGs, the Paris Declaration on Aid Effectiveness and the SDGs, which has consequently resulted in emerging challenges and demands in evaluation. As Conlin and Stirrat (2008) asserted a decade ago:

The situation facing development evaluators today is vastly more complex than it was 15 or even 10 years ago. Shifts in the nature of aid, coupled with changing relationships between donors and recipients, have made it much more difficult to produce neat and firmly demarcated forms of evaluation. (p.197)

This observation remains valid. Conventional models rooted in positivist principles and derived from project logical frameworks (logical framework is a widely adopted project management tool) remain prominent and dominant among programme implementers and evaluators with strong support from donor agencies. However, they are neither very adaptive to the fluid and complex development context, nor well able to address the increasingly recognized diversity of stakeholders and actors involved in the development process (ibid.). Under this circumstance, some key evaluation challenges, such as the dichotomy between learning and accountability and the attribution issue, remain to be adequately addressed. The following sections unpack each of these evaluation challenges in turn to further identify the gaps and needs so that potential solutions can be possibly sought.

### **2.3.1 Challenges in reconciling learning and accountability**

The first key challenge lies in the tension between the two major goals of evaluation: learning and accountability. This challenge has been long identified but remains unresolved to date, and may be even potentially exacerbated within the accountability-focused evaluation and management system in the current development evaluation context (Reinertsen et al., 2017). The two are considered to be in tension as ‘they operationalize the evaluation function using different data, methodologies and incentives’ (Armytage, 2011, p.263). Metaphor such as ‘a warring couple’ (Adams, 2007) has been used to depict and illustrate this complex relationship and the challenge in realizing both learning and accountability in evaluation. More specifically, as observed by Cracknell (2000), learning- and accountability-oriented

evaluation differ in many aspects, such as basic aim, potential users, evaluator role, timing of evaluation, and data and feedback needs. These views seem to mainly focus on accountability to donors and funders of development programmes, but accountability to aid programme beneficiaries may also be relevant. Other researchers such as Reinertsen et al. (2017) argue that direct contradictions may arise due to diverging accountability and learning needs during evaluation. For example, accountability principles of ‘critical distance and independence’ in the evaluation process contradict with critical enabling factors for learning – ‘building and sustaining internal engagement’ of actors involved for the evaluation (ibid., pp.14-15).

In addition to incompatibility and divergence identified in the dual goals of evaluation, there are also various contentions that oppose an either/or situation in terms of learning and accountability. Deliberately separating the two may create a false dichotomy, and learning can actually be built around accountability rather than being opposed to it (OECD, 2001). For example, the World Bank (cited in OECD, 2001, p.17) suggests that accountability should create an ‘incentive framework for learning’. Moreover, there has been a tendency to shift from an accountability or audit focus in development evaluation theory to an effectiveness or learning focus (Armytage, 2011; Lennie and Tacchi, 2014). And more recently, there has been a growing demand for learning (OECD, 2016) and a call to place learning needs at the centre so that development programmes can be adaptive<sup>2</sup> (Valters, et al., 2016). It is therefore key to explore possibilities to reconcile learning and accountability (more detailed and nuanced discussion in Section 2.4) and the critical role of evaluation approaches in this regard should be noted.

---

<sup>2</sup> In the sense of responding to changes in circumstances or to new information about the efficacy of the programme (Valters, et al., 2016, p.5).

### 2.3.2 Challenges in addressing attribution

Regardless of the primary focus of evaluation, attribution remains one of its basic practical challenges (Crawford et al., 2004; Mayne, 2001; Smutylo, 2001; Solmeyer and Constance, 2015). Attribution problems intensified under the ‘focusing on results’ agenda initiated in the wake of the MDGs and shift in emphasis from input to outcome (White, 2010). The demand to demonstrate development impact led to issues in attributing results and claiming credit during evaluation, due to increasing acknowledgement of the complexity of the development context and multiple influencing factors involved (Conlin and Stirrat, 2008; White, 2010).

As a central issue in evaluation (Patton, 2012), attribution refers to ‘identifying causality between the intervention and the changes observed’ (Crawford et al., 2004, p.176). Such causality involves building logical and causal links and relationships between activities and outputs, as well as between the intervention and its outcomes or impacts<sup>3</sup>.

Establishing attribution at output level is relatively feasible and straightforward, but issues can quickly arise when seeking to make such a link and claiming credit for accomplishment at outcome or impact level (Mayne, 2001). As Jones (2006, p.8) points out, ‘The “attribution problem”, the fact that in international development it is nearly impossible to prove the extent to which a development outcome as been caused by an individual agency or programme, has long been used to justify a focus on achieving predetermined outputs ahead of driving for real long-term change’. Such a challenge in proving attribution is inevitable as development assistance evolves, and

---

<sup>3</sup> The terms activity input, output, outcome and impact mentioned here generally follow the conventional definition in logframe terms, as listed in Table 3.3 in Chapter 3 (Bakewell and Garbutt, 2005).

development impacts in particular can rarely be achieved by the work of a single programme or actor (*ibid.*).

Moreover, the attempt to simply demonstrate and attribute credit for development outcomes or impacts may severely limit one's potential to understand how and why outcomes/impacts occur. Merely addressing attribution may not enable a proper identification and examination of the contributions of other actors and contextual factors, even though they jointly produce development results (Earl et al., 2001). Additionally, some aspects of development such as research-oriented activities usually require a considerable period of time before intended outcomes occur, and, particularly if this is beyond the project cycle, it would then be unrealistic to demonstrate and claim attribution for impact (Earl, et al., 2001; Mayne, 2001).

From a methodological viewpoint, there are some practical challenges (e.g. limited time and financial resources and lack of quality data) in implementing those approaches recognized as more capable of achieving attribution at impact level – more positivist-oriented and statistically strong methods such as experimental designs (Armytage, 2011). Those who consequently choose to rely primarily on a seemingly simpler method such as the LFA tend to mainly assess output. They may often have an underlying theory of change (a theory of how and why the initiative works) linking results at output level to impact (Weiss, 1995), but consider that it is too costly or too complex to actually try and seek data to demonstrate such a connection, and instead simply assume it.

More recently, as Gates and Dyson (2017) observe, there has seen an increasing emphasis on exploring causal process and relationships between development

programmes and their results/effects. With more evaluations commissioned for this purpose, ‘how evaluators warrant causal claims has recently come under considerable scrutiny and debate’ (ibid., p.30). This suggests increasing attention and concern about the key role of evaluation design and methodologies in addressing the attribution issue. There is also growing interest in understanding the key elements and mechanisms functioning in the programme which effect changes and results to make explicit how and why a programme works (Solmeyer and Constance, 2015). However, evaluations are considered to be ‘rarely designed to rigorously examine such questions’ (ibid., p.470). These evaluation demands necessitate evaluation approaches with explanatory power to discern the intermediary mechanisms/links connecting outputs and outcomes or goal<sup>4</sup>, so as to unpack the change process and answer relevant evaluation questions.

Tackling the attribution issue means recognising the multiplicity of factors which influence results, and of actors who may have been involved in producing them, which in turn calls for an evaluation approach which explicitly seeks to clarify roles of different actors. It also entails putting in place comprehensive and adaptive evaluation approaches with a degree of explanatory power to embed such consideration, reflect the complexity and unpack the causal links during evaluation. In this way, attribution at various levels can possibly be addressed in a stratified and feasible way under real-world constraints in a development context.

The above overview of key aspects of development evaluation (objectives, methodological and epistemological evolution, and existing and emerging challenges and demands) provides the context in which this research project is grounded. By

---

<sup>4</sup> There will be a more nuanced differentiation of outcome and goal with possible variations later in this thesis in relation to clearly identified corresponding groups of actors.



outlining this broader evaluation context and contemporary evaluation challenges and demands in particular, it is indicated that, as concluded by Lennie and Tacchi (2014, p.13), ‘a broader range of evaluation approaches and methodologies that can better address the complex challenges and issues in the evaluation of development initiatives’ are needed.

## **2.4 Overview of relevant concepts**

In the following sections, I review some key concepts drawn from the above literature review especially these identified evaluation challenges and demands – learning, accountability, explanatory power of evaluation approaches (in terms of explaining and unpacking the mechanisms linking outputs with outcomes or goal) and actor-centred perspective (given its potential value added as reviewed earlier). These concepts form the basic building blocks and components for my conceptual framework that can be used to assess to what extent an evaluation approach can help address identified evaluation challenges and demands.

### **2.4.1 Learning in evaluation**

In Section 2.1, the OECD definition of development evaluation explicitly points out that learning is one of the key goals of evaluation with the ultimate aim of informing the decision-making process. In the literature, there are diverse definitions and interpretations of learning from various starting points and perspectives. As Minsky (1988, p.120) summarizes: ‘The problem is that we use the single word “learning” to cover too diverse a society of ideas’. This section mainly focuses on the presence of the concept of learning in the context of development evaluation.

Kolb’s experiential learning notion is widely adopted – ‘Learning is the process

whereby knowledge is created through the transformation of experience’ (Kolb, 1984, p.38). Kolb considers learning to be a cyclical process of experiencing, reflecting, conceptualizing and experimenting with newly gained knowledge (ibid.). Toulemonde (2016, p.16) provides a very succinct definition of learning which highly summaries Kolb’s four-stage learning cycle – ‘Acquiring and using knowledge’. In contrast, Meadows et al. (2005) define learning in a more detailed manner:

Learning means the willingness to go slowly, to try things out, and to collect information about the effects of actions, including the crucial but not always welcome information that the action is not working. One can’t learn without making mistakes, telling the truth about them, and moving on. Learning means exploring a new path with vigor and courage, being open to other people’s explorations or other paths, and being willing to switch paths if one is found that leads more directly to the goal. (p. 7)

This definition considers mindset and action of learners, information type (both positive and undesired), pathway and means of learning, which corresponds to the observation of Preskill (2008, p.129) that ‘learning is inextricably linked to change’ – changes in thinking, behavior, and/or beliefs.

These generic definitions of learning are readily applicable to the evaluation context. There is a lack of a clear definition of learning in evaluation and the concept is subject to various interpretations and variations by different individuals and agencies (Kogen, 2017). For example, Guijt (2010, p.281) defines learning, in a development programme setting, as ‘the process of continual reflection about visions, strategies, actions and contexts that enable continual readjustments’. In addition to embedding the basic notions of learning such as knowledge processing and utilization, this definition explicitly brings context into consideration. United States Agency for International Development (USAID, 2016, p.4) defines learning as ‘systematically

generat[ing] knowledge about the magnitude and determinants of project performance, permitting those who design and implement [projects].....to refine designs and introduce improvements into future efforts'. This version exemplifies some core purposes of learning in evaluation such as making practical improvements and strategic adjustments (Guijt, 2008), but may also reflect what Levinthal and March (1993, p.101) identified as 'myopia' of learning – a 'tendency to ignore the long run' and a 'tendency to ignore the larger picture'. Kogen (2017, p.103), after more than two decades, echoes this view by pointing out: 'Learning should mean much more than refining a single project'.

To further unpack the learning concept, OECD (2001, p.18) outlines four stages in the learning process: seeking new knowledge on results, 'validating and verifying' them, transferring to other cases, and then 'codifying them into guidelines'. However, transferring knowledge seems more challenging in reality as it requires careful consideration and study of contextual factors. Specifically, in evaluation context, as indicated in OECD (2002) definition of development evaluation, the achievement of learning involves acquiring credible and useful information on effects and results of interventions, processing the information and drawing lessons. And finally and most critically, these 'lessons need to be absorbed and acted upon' (e.g. in the decision-making process) before we can really realize learning in evaluation (Toulemonde, 2016, p.16).

There are diverse ways of categorizing learning such as by levels of learning (e.g. individual or collective); by purposes and loops of learning such as single- and double-loop learning (Argyris and Schön, 1978). Single-loop learning mainly looks at whether things are done well through the comparison between actual and expected

results and involves necessary adjustment of actions without looking into underlying assumptions; Double-loop learning questions ‘are we doing the right things’ and looks into the underlying assumptions and values and makes modifications if necessary (Guijt, 2008; 2010, p.282).

As discussed in Section 2.3.1, there has been an increasing stress on how to better achieve learning in evaluation. There remains a gap of knowledge in this regard due to a lack of consensus on the definition of learning and diverse ways of applying the concept. For example, there is little insight on how to define learning and limited information on anticipated learning product and required data in the evaluation strategy documentation of DFID (2014) (Kogen, 2017).

Toulemonde (2016, pp.16-17) observes that learning-oriented evaluations ‘tend to focus on challenging knowledge gaps, target policy-making windows, cover a narrow scope, use a few evaluation criteria or just one, seek generalizable lessons, explain the reasons of successes as well as failures, and establish close links with field level agents’. This is a comprehensive list of learning features but rather challenging to achieve in one evaluation given current status of learning and time and resource constraints in real world. From an operational point of view, USAID (2016) summarizes that learning in evaluation requires:

careful selection of evaluation questions to test fundamental assumptions underlying strategies and project designs; methods that generate findings that are internally and externally valid (including clustering evaluations around priority thematic questions); and systems to share findings widely and facilitate integration of the evaluation conclusions to recommendations into decision-making. (p.4)

This summary explicitly suggests the methodological requirement in achieving

learning apart from sharing some key elements with the observation of Toulemonde (2016). Furthermore, Preskill (2008) argues that taking a learning perspective in evaluation entails mindset changes in terms of revisiting and rethinking underlying assumptions and beliefs, and updating the understanding of the intervention concerned.

However, in addition to the confusion around what is the best way and precise content to learn and how to extract useful lessons (Kogen, 2017), there are considerable practical obstacles and challenges for achieving quality learning. At organizational level, lack of openness to accept errors in organizational culture and rigid accountability-based administrative system may create barriers for learning incentives, process and outcome (Guijt, 2010; OECD, 2001; Reinertsen et al., 2017). Learning uptake usually involves a ‘long chain of events’ (Toulemonde, 2016, p.16) – continuous reflection, collection and analysis of evidence and then action upon lessons learnt – in a relatively long time (Adams, 2007; Guijt, 2010). However, staff at operational level may often have limited time, capacity and skills for fulfilling all these learning activities (ibid.) and they tend to keep to their comfort zone and familiar approaches in their operations, this might create ‘tunnel vision’ (OECD, 2001, p.21).

In addition, and most importantly, corresponding to previously identified evaluation demand in Section 2.3.2, Kogen (2017, pp.101-102) emphasizes that ‘learning why particular programs work in particular contexts is what contributes to improving policy, not simply learning whether programs worked’. However, this question remains unaddressed in many accountability-based evaluations (ibid.).

There are no immediate solutions to some of these barriers and challenges to learning such as organizational culture, administrative system and staff mentality, but limitations in capacity and skills may be possibly addressed in a relatively short term. Better understanding of the relationship between learning and accountability in evaluation is also critical in exploring possibilities to achieve quality learning. The next section proceeds to the review of accountability in evaluation.

#### **2.4.2 Accountability in evaluation**

Accountability, one of the key goals of evaluation which often gets prioritized in practice (Regeer et al., 2016), is regarded as a ‘complex and somewhat ambiguous construct’ (Ebrahim 2005, p.60). Indeed, there are diverse ways of conceptualizing and interpreting accountability. For example, OECD (2002) provides two definitions of accountability. The first is more general and outlines some key aspects of accountability (in a contractual relationship) such as obligation, compliance, and reporting on performance results. The second, as cited below, is particularly about accountability in a development setting and differentiates connotations for different development actors including evaluators.

Accountability in development may refer to the obligations of partners to act according to clearly defined responsibilities, roles and performance expectations, often with respect to the prudent use of resources. For evaluators, it connotes the responsibility to provide accurate, fair and credible monitoring reports and performance assessments. For public sector managers and policy-makers, accountability is to taxpayers/citizens. (OECD, 2002, p.15)

In the context of development evaluation, accountability ‘has customarily become associated with the judgment of whether a program or a policy has achieved its objectives’ (Lehtonen, 2005, p.175). More specifically, as Toulemonde (2016, p.16) observes, accountability-oriented evaluation tends to ‘focus on major expenditures,

follow programming cycles, cover a wide scope, include the whole range of evaluation criteria, favor internal validity, focus on potential risks and problems, and highlight independence'. These two perspectives on accountability in evaluation both imply a major focus on financial accountability.

Accountability is complex and 'multi-layered' especially in the development field (Valters et al., 2016, p.19). This has been increasingly emphasized over recent years and the necessity of considering multiple forms of accountability has been recognized (Lennie and Tacchi, 2014; Valters et al., 2016; Van Ongevalle et al., 2014). There exist different perspectives and approaches in differentiating and categorizing these forms and some argue that 'there are as many types of accountability as there are distinct relationships among people and organizations' (Ebrahim, 2005, p.60).

For example, Schillemans (2008, pp.175-176) differentiates 'vertical' accountability – 'traditional forms of accountability' in which 'a subordinate usually reports to a superior' – with 'horizontal' accountability which addresses 'peers, equals, stakeholders or concerns outside of the hierarchical relationship between central government and executive agency'. Accountability is relational (Ebrahim, 2005; Guijt, 2010) and relative; the reference point in clarifying whose accountability is therefore critical. Taking implementing organizations as the reference point, the following three forms of accountability are often used in the development context (Ebrahim, 2005, p.60):

- 1) Upward accountability (e.g. to donor agencies and governments) mainly focuses on financial aspect – 'spending of designated moneys for designated purposes' (ibid., p.60) – it is sometimes used interchangeably with financial accountability (Regeer et al., 2016).

- 2) Downward accountability mainly refers to relationships with service recipients or beneficiaries.
- 3) Internal accountability relates to implementing organisations themselves and their responsibilities internally to organizational mission and staff members both at management and field level.

This approach of differentiating accountability has some overlap with that of Schillemans (2008). For example, vertical accountability may overlap with upward one and horizontal accountability with internal one. In this system of accountability, upward or financial accountability tends to be the common conceptualization (Regeer et al., 2016) and often gets prioritized over other types of accountability in practice (Ebrahim, 2005). This risks privileging one type of accountability and focusing on mainly some relevant actors such as donors but neglecting others such as implementing agencies and beneficiaries (ibid.). This has certain implications for evaluation.

Evaluation driven by upward accountability helps assess ‘performance, or effectiveness, as well as efficiency of organizations, programmes or projects by gauging resource use in relation to services provided and impacts achieved’ (Regeer et al., 2016, p.12). However, it also tends to promote preset quantitative performance indicators and relatively simple and measurable objectives (Kogen, 2017), especially under the current results-based development agenda. This may result in myopia in terms of neglecting long-term results and parameters as well as larger picture of social change (Ebrahim, 2005; Kogen, 2017). In addition to these identified problems of (upward) accountability-oriented evaluation, there exist certain challenges and obstacles to fully achieve accountability in evaluation as programme results are often gathered from arguable attribution or causal claims and involve responsibilities from



multiple stakeholders (Toulemonde, 2016), as already discussed in Section 2.3.2. The accountability function of evaluation therefore rests heavily on dealing adequately with questions of attribution.

Given above-mentioned challenges and issues in relation to accountability in evaluation, Kogen (2017) argues that: 1) re-conceptualization of accountability is needed to look at more fundamental and broader range of issues of development (e.g. promoting social change) in addition to financial aspects (e.g. accounting mechanisms); 2) accountability should be treated as a secondary objective of evaluation and as a means to an end – learning about how to effectively improve development should be prioritized. Although there have been some changes in terms of increasing attention to internal and downward accountability in evaluation in recent years (Lennie and Tacchi, 2014), a better and more dynamic understanding and consideration of various levels of accountability should be enhanced from an actor-oriented perspective (Van Ongevalle et al., 2014).

These proposals not only entail changes in mindset but also in evaluation methodologies. Adaptive evaluation approaches need to be explored to help fulfil these suggested changes. Some researchers such as Regeer et al. (2016) propose that learning-oriented evaluation approaches using, for example, participatory methods, can possibly help include and strengthen downward accountability, while Kogen (2017) point out a lack of sufficient evidence on how participatory approaches get used and perform in practice.

### **2.4.3 Learning and accountability in evaluation**

As reviewed in Section 2.3.1, there are divergent views about the relationship

between learning and accountability in evaluation, and they mostly focus on upward/financial accountability. As reviewed in the previous section, upward accountability is often prioritized in evaluation, generating a tension rather than a complementarity with learning. Given the importance of both goals in helping ensure evaluation achieves better development results, the divide between the two needs to be bridged. Some development researchers and practitioners have sought to explore possibilities to reconcile learning and accountability.

For example, Ebrahim (2005) suggests that upward accountability and learning can share a space but it requires changes in staff perception about evaluation and learning, organizational culture (e.g. more open to errors and less blame), organizational capacities, internal reporting structure and information system (e.g. guided by internal rather than upward accountability). These suggestions are made from the starting point of implementing organizations and most of them seem to be proposed to address learning barriers (as identified in Section 2.4.1). They need concerted efforts among organizational staff at various levels over long time. Thus, it is rather challenging to realize these proposals in practice. Guijt (2010) shares similar view that learning and accountability can be reconciled and proposed some more general principles such as the necessity of clarifying the two concepts, necessarily merging accountability and learning needs and activities, better understanding of the nature of anticipated changes and power dynamics among relevant actors. The importance of appropriate evaluation methodologies was also pointed out (*ibid.*). Similarly, Reinertsen et al. (2017) argue that methodological rigour may contribute to reconciling diverging needs of learning and accountability in evaluation.

Some recent literature reveal that internal and downward accountability could be

possibly achieved by, for example, learning-oriented evaluation methodologies (Regeer et al., 2016), but there still remains question about ‘how to align evaluation for upwards accountability with evaluation for learning towards system change’ (ibid., p.13). Furthermore, from methodological perspective, Regeer et al. (2016) point out that separate approaches are often adopted to address learning and accountability in evaluation at different phases of the intervention, this may result in further divide between the two evaluation goals. One unified evaluation approach may be an alternative worth exploring to help resolve the tension between the two and address them concurrently (ibid.).

#### **2.4.4 Actor-centred perspective**

Essentially, ‘development is accomplished by, and for, people’ (Earl et al., 2001, p.2). The central role played by ‘human action and consciousness’ (Long, 2001, p.13) needs to be recognized and reflected in the understanding of development process. Taking an actor-centred perspective or approach helps serve such purpose. This type of approach gained popularity in sociology and anthropology in the late 1960s and early 1970s and has been gradually adopted in the development context<sup>5</sup> (ibid.).

As briefly discussed in Section 2.2.1, development intervention can be seen as a transforming process of internalizing external factors and resources. In this process, an actor-centred approach recognizes social actors as active participants rather than passive recipients of development interventions and considers different implications for and responses of various groups of actors (ibid.). It thus could offer valuable insights into and understanding of the social construction and reconstruction process.

---

<sup>5</sup> According to Long (2001), the usage of actor-oriented approaches was less discussed in development literature until the twentieth century.

More specifically, as Van Ongevalle et al. (2012) observe, during a transformational process of internalizing external interventions:

Multiple actors, relationships and perspectives involved in complex-change processes mean that the achievement and sustainability of programme results often depend on changes in behaviour or practices by multiple intermediate actors who in turn interact (directly or indirectly) with programme beneficiaries. As a result, a programme's resources and efforts move through a chain or network of intermediate actors before an eventual effect may be felt by the target group. (Van Ongevalle et al., 2012, p.6)

This explicit actor-oriented perspective on change process has certain implications for evaluation. First, this implies that the actor component should be made more evident in evaluation and various groups of actors along the actors chain should be differentiated, for example, intermediate actors and programme beneficiaries, as identified in above citation; Second, the interactions among a multitude of actors need to be considered (Davies, 2005), these interactions may take different forms such as 'collaboration, negotiation, dialogue, influencing, lobbying and conflict' (Van Ongevalle et al., 2014, p.449); Third, in the process of interactions among various actors, different features and characteristics of these actors need to be considered. There might be 'messy partnerships' (Guijt, 2008, p.20) formed due to differences among actors in terms of 'governance structure, culture, mandate, capacities, priorities and commitment to collective efforts' (ibid.).

An actor-centred perspective to evaluation has positive implications for its two key goals – learning and accountability. As Guijt (2010, p.289) argues, 'Processes of learning and accountability are essentially about bringing people and their perspective together to make sense of information and value performance in order to be able to respond'. This remains challenging in practice, as discussed in respective sections on learning and accountability. Specifically from an actor-oriented perspective, in terms

of accountability, as mentioned in Section 2.4.2, there exist obstacles to fully achieve this goal in evaluation as it may be difficult to identify accountable actors due to shared responsibilities among stakeholders especially in a multi-actor and highly interactive context (Van Der Meer and Edelenbos, 2006). In this regard, taking an actor-centred approach in evaluation may help bring clarity by identifying different groups of actors, and then differentiating and relating these actors respectively to their corresponding domains of responsibilities, roles and contributions to the intervention – which helps make more credible causal claims to respond to the attribution issue. This approach also provides possibility to address accountability at multiple levels (e.g. internal and downward accountability) once actors in relation to each level are explicitly identified.

With regard to learning, as discussed in Section 2.4.1, it involves a long chain of events in terms of processing information and acquiring knowledge to then effect behavioural changes of actors involved (Ebrahim, 2005). Different perspectives, interests and capacities of various actors could all possibly influence this learning process and its outcome. With an actor-centred perspective in evaluation, explicit identification and differentiation of actors involved at various stages of the intervention enables a more nuanced understanding of the characteristics and needs of different actors. Learning activities and strategies can be therefore better planned and tailored and capacity and skill needs for effective learning can be more easily addressed. More critically, as Van Ongevalle et al. (2014) note, an actor-centred approach to evaluation could help draw attention to those easily neglected results in relation to behavioural, relationship and mindset changes among actors, as well as those unexpected and less easily measurable effects. In this way, it helps to better learn about what and how actual changes occur in relation to corresponding actors.

They further argue that actor-focused monitoring and evaluation practice is not ‘just an interesting complement to more mainstream linear planning logic’, it is ‘an essential component of learning-centred programme management, particularly in contexts of complex change’ (ibid., p.462).

Despite these potential value added of an actor-centred perspective to evaluation, there remains a key question to be addressed: how can evaluation (with an actor-centred perspective) help ‘identify accountable actors and their domains of responsibility, and at the same time facilitate learning by different actors (with their diverse perspectives) and realize collective action?’ (Van Der Meer and Edelenbos, 2006, p.204). Furthermore, some of the widely adopted evaluation tools such as logical frameworks tend to abstract broader range of human actors but mainly focus on one particular type – the project implementation team (Van Ongevalle and Fonteneau, 2014). An actor-centred perspective in evaluation remains to be better recognized and embedded in practice. On the other hand, when adopting an actor-oriented approach, the potential risk of excessive or over reliance on explanations based on individual interests and motivations should be avoided (Long, 2001).

#### **2.4.5 Explanatory power – unpacking the mechanisms linking outputs and outcomes or goal**

As pointed out in Section 2.3.2, with growing stress on exploring causal process and relationships between development programmes and their attributable results and increasing interest in unpacking the functioning mechanisms of a programme to understand how and why the programme works, there has been considerable concern about the role of evaluation design and methodologies. These recent tendencies and concern imply the necessity of seeking adaptive evaluation approaches with

explanatory power to unpack the intermediary mechanisms/links connecting outputs and outcomes or goal.

In a general sense, as Colombo (2017, p.503) observes, ‘While all explanations answer some why- or how-question, significant variation is observed across contexts in what is accepted as an explanation, in what type of explanatory information is sought, and in what norms are assumed to govern good explanation.’ Explanatory power is variously defined, and its various usages will give rise to different ways of determining or assessing it, but in essence, ‘powerful explanations provide information about credible causal relationships’ (Colombo et al., 2017, p.2). This observation, while coming from the field of psychology, is relevant and applicable to the context of development evaluation. Explanations involved in this thesis aim to help answer how and why a programme works through unpacking the mechanisms connecting outputs and outcomes or goal. In terms of mechanism, it can be conceptualized differently according to the ‘scope of the intended explanation’ (Dalkin et al., 2015, p.2). This thesis mainly draws on a realist account of mechanism in evaluation. As briefly reviewed in Section 2.2.2, Pawson and Tilley (1997, p.66) point out that a programme mechanism focus needs to unpack what particular aspects or components of a programme make it work. They further assert that programme mechanisms in practice should exhibit the following three features and functions:

- 1) Reflect the embeddedness of the programme in a stratified social reality.
- 2) Indicate how both macro and micro processes construct the programme.
- 3) Demonstrate how programme outcomes come about in relation to stakeholders’ choices (reasoning) and capacity (resources) to operationalize them (ibid.).

Identifying mechanisms by probing beneath surface phenomena can thus considerably

aid deep understanding and explain how programme effects and outcomes are generated, as well as relationships between causes and effects. The above-mentioned features also outline key operational-level elements of programme mechanisms such as resources and reasoning (Dalkin et al., 2015). Or in another words, mechanisms are ‘a combination of resources offered by the social programme under study and stakeholders’ reasoning in response’ (ibid., p.3). This corresponds to the observation of Weiss (1997) that mechanisms are the responses (of stakeholders) triggered by the programme service.

These viewpoints about mechanism take an actor perspective, as reviewed in the previous section, when external factors and resources enter the existing lifeworlds of those actors involved and become internalized, different interest groups or actors may then respond in various ways. To better understand this process, it is necessary to identify different groups of actors along the actors chain and differentiate their respective reasoning and responses to the intervention so as to have more explanatory value.

In terms of intermediary mechanisms connecting outputs and outcomes or goal, as discussed in the previous section, dominant project monitoring and evaluation tools such as logical frameworks tend to abstract human actors – mainly focusing on one type of actors (project implementation team) but easily neglecting intermediary actors (e.g. partner organizations working directly with the project implementation team) and beneficiaries. This may then result in more attention being paid to outputs of project activities than to effects such as outcomes or goal along the results chain (Van Ongevalle and Fonteneau, 2014). There may therefore exist a ‘missing’ or under-explored link between outputs and outcomes or goal. With the shift of evaluation



focus from outputs to outcomes under the results-based development agenda and more emphasis on exploring causal relationships between development interventions and their results, unpacking the intermediary mechanisms linking outputs and outcomes or goal helps fulfil such demands and strengthens the explanatory power of evaluation. Taking a realist approach of conceptualization (according to which programme mechanisms combine resources, stakeholder reasoning and responses, as discussed earlier in this section), this intermediary mechanism may then involve unpacking changes in reasoning, responses and behaviours of intermediary level actors with whom the programme implementation team interacts directly.

Adding explanatory power to evaluations by looking into the programme mechanisms at work not only helps provide grounds for making causal claims and telling performance stories of a programme, but also facilitates learning between programmes. As Astbury and Leeuw (2010) observe, there may be common underlying mechanisms at work in different development interventions, and accumulation of knowledge about mechanisms could help better inform future policy or programme design.

The importance of seeking attribution and clarifying causal relationships in evaluation notwithstanding, it remains an ongoing challenge for evaluators to warrant causal claims (Gates and Dyson, 2017). Appropriate evaluation approaches need to be pursued that have greater explanatory power to uncover the mechanism or links among different levels of results in the results chain.

## **2.5 Conceptual framework**

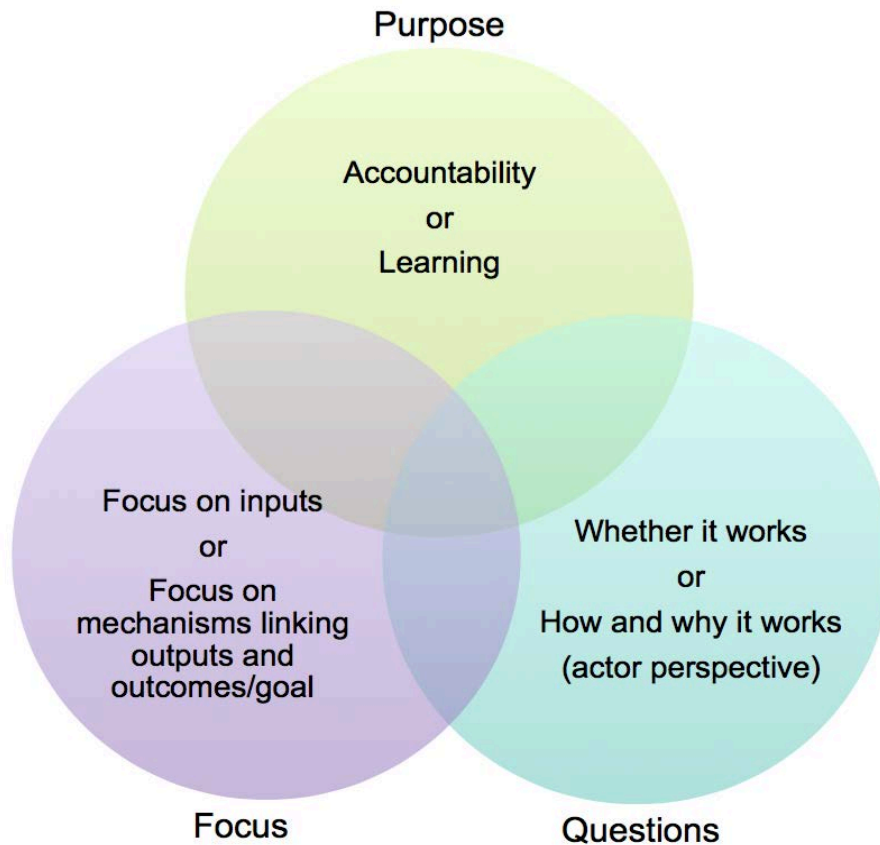
In this section, I formulate the conceptual framework for this research project.

Building on the four reviewed concepts and perspectives – learning, accountability, an actor perspective, and a mechanism-based explanation unpacking the links between outputs and outcomes or goal – this conceptual framework (as shown below in Figure 2.1) connects these components by locating them within key areas of concern in evaluation (purpose, focus and questions). My research focus is situated in the domain where these key areas converge.

Figure 2.1 captures this visually, setting out in three overlapping circles the range of purposes, focuses and questions addressed in evaluations. Within each circle (Purpose, Focus, and Question) are summed up the respective tensions or evaluation challenges. For reasons of diagrammatic simplicity, these are expressed here as dualisms or as two extremes; it needs to be recognised that in practice they are not always in opposition or an either/or, dualistic, relationship, but they are shown as such so as to represent the real tensions that tend to arise when both are pursued concurrently in any given evaluation undertaking, as discussed in the foregoing sections. The green circle at the top shows two major purposes and objectives of development evaluation – learning and accountability. The right-hand blue circle shows key questions pursued in evaluation, reflecting the increasing emphasis on questions such as how and why a programme works in addition to whether it works, and the potential value added of an actor perspective in helping address these questions. The left-hand purple circle illustrates how some evaluations focus on inputs whereas others focus on ‘mechanisms’, unpacking and explaining the links between outputs and outcomes or goal, to help supplement understanding of the connections between inputs and outputs and make the evaluation story more complete along the results chain. My research then focuses on where these three key areas of concern overlap – potentially exploring responses to some key issues connecting these

three evaluation aspects.

**Figure 2.1 Conceptual framework**



In this overlap area, these involved elements are closely linked and may become potential enabling factors for one another. For example, a focus on intermediary mechanism linking outputs and outcomes or goal strengthens the explanatory power of evaluation and helps answer questions of how and why a programme works, thus also fulfilling an important learning function; an actor perspective could help facilitate and operationalize this process by making micro-level mechanism of change evident through further identification of relevant actors and their interactions (e.g. reasoning process and behavioural pattern of individual groups of actors). They then work in synergy and jointly contribute to learning and accountability (in addition to their respective value and function in helping address the two key evaluation objectives, as

reviewed previously in this chapter) and potentially create a shared space for the two in evaluation.

In this chapter, from the review of broader evaluation literature and a closer examination of key concepts and themes drawn from this review, a need for further exploring innovative evaluation approaches to address key evaluation challenges and demands has been identified. This above conceptual framework provides the parameters to assess the evaluation approaches to be explored in this thesis — the extent to which the approach under study helps reconcile the tension between learning and accountability and elucidate mechanisms linking outputs and outcomes or goal by incorporating a more actor-oriented perspective.

## **Chapter 3 Critical review of existing approaches in the evaluation toolkit and exploration of potential alternative**

Following the identification of the need for pursuing innovative approaches to respond to contemporary evaluation challenges and demands, and oriented by the conceptual framework constructed in the previous chapter, I focus on evaluation methodologies in this chapter and conduct a critical review and secondary analysis of selected approaches currently available in the evaluation toolkit. This is structured with a view to elaborating on the position set out in the previous chapter that there are some methodological gaps in the evaluation toolkit in respect of certain tensions and dilemmas prevalent in contemporary evaluation debates and practice. Section 3.1 provides a general overview and analysis of some approaches in the evaluation toolkit. Section 3.2 focuses on the Logical Framework Approach and Outcome Mapping as two key, dominant, evaluation methodologies, and critically reviews their role in evaluation in a detailed manner. Section 3.3 introduces a synthesis model of the LFA and OM which has featured very little in evaluation literature and debates, as an alternative evaluation approach which potentially adds value to current dominant approaches in its scope to deal with evaluation challenges and demands, as reflected in the conceptual framework.

### **3.1 Overview of selected approaches in the evaluation toolkit**

The previous chapter pointed to the need for appropriate and effective approaches to help address certain contemporary evaluation challenges, in particular to reconcile the tension between learning and accountability, and to offer explanatory power by elucidating mechanisms linking outputs and outcomes or goals, and by adding an actor perspective. This section reviews prevalent approaches in the evaluation toolkit to see how they respond to these challenges and demands, with a view to further

identifying methodological gap and potential solutions.

Various typologies have been used in the evaluation literature to classify evaluation approaches. For example, Stern et al. (2012) observe six types of evaluation approaches according to their respective means of making causal claims, including experimental (using counterfactuals, as in Randomized Controlled Trials), statistical (correlation between cause and effects), theory-based (evaluation based on or guided by the theory of how the initiative or intervention brings about results – showing causal process, supporting and contextual factors and mechanisms at work), case-based (analysis across or within cases, using grounded theory and etc), participatory (involving validation by key participating actors) and synthesis studies (aggregation of different perspectives or approaches). Hansen (2005) provides a different typology with another set of six evaluation models, as shown in below Table 3.1, according to different focuses, questions, and criteria to be pursued during evaluation.

**Table 3.1: A typology of evaluation models**

<b>Evaluation models</b>	<b>Description and focus</b>	<b>Questions</b>	<b>Criteria for Evaluation</b>
<b>Results models</b>			
1) Goal-attainment model	Results assessed according to preset goals	To what degree has the goal(s) been realized?	Derived from goal(s)
2) Effects model	Involving all results of the intervention	What effects can be uncovered?	Open, all effects should be uncovered
<b>Explanatory process model</b>	Focusing on ongoing processes or actions	Is the level of activity satisfactory? Are there implementation problems?	Performance is analysed from idea to decision and implementation, and to reactions of the addressees
<b>System model</b>	Analysing input, structure, process and outcome	How has performance functioned as a whole?	Realized input, process, structure and outcome assessed either in relation to objectives in same dimensions or comparatively
<b>Economic model</b>	Evaluation object considered to be a ‘black box’ and result assessment involving expenses	Is productivity satisfactory? Is effectiveness satisfactory? Is utility satisfactory?	Output measured in relation to expenses Effect measured in relation to expenses Utility measured in relation to expenses
<b>Actor model</b>	Reliance on actors’		

1) Client-oriented model	assessment criteria	Are clients satisfied?	Formulated by clients
2) Stakeholder model		Are stakeholders satisfied?	Formulated by stakeholders
3) Peer review model		Is professional quality in order?	Formulated by peers
<b>Programme theory/theory-based evaluation model</b>	Focusing on the validity of underlying theory	What works for whom in what context? Is it possible to ascertain errors in programme theory?	Programme theory is reconstructed and assessed via empirical analysis

(Adapted from: Hansen, 2005, pp.449-450)

The above-mentioned two typologies have different starting points but also overlap to some extent – both explicitly mention actor-oriented and theory-based evaluation model. Situating these typologies under the broader methodological evolution of evaluation as reviewed in the previous chapter, the one from Hansen (2005) in particular offers a more specific differentiation of evaluation models at an operational level. However, these two typologies do not seem to explicitly reflect aspects on the practical usage and function of various models in evaluation.

Eight commonly used tools and approaches are selected and analysed in this section which can all be considered part of the evaluation ‘toolkit’ – RCT, Theory of Change, Contribution Analysis, Political Economy Analysis, Value for Money, Cost-Benefit Analysis, LFA, and OM. The rationale for selecting these is to exhibit the variety of choices in the evaluation toolkit in terms of evaluative features – which I distinguish as: technocratic or actor-centred; political or economic; function (output, outcome, impact or process evaluation); diagnostic or facilitating role; and applicability. This further differentiation attempts to build additional dimensions onto the typologies set out above, so as to better capture the applicability aspect. The analysis of selected approaches, presented in below Table 3.2, incorporates aspects of both Stern et al.’s and Hansen’s typologies as well as these evaluative features, and seeks to reveal matches and gaps between existing tools and key evaluation challenges and demands, as reflected in the conceptual framework, so that unaddressed needs can be identified

and relevant effort made.

**Table 3.2: A critical review and analysis of eight selected approaches**

Selected approach	Key evaluation features and epistemological stance	Strengths	Weaknesses	Evaluation model	Application coverage/ When should this approach be used
<b>Randomized Controlled Trial</b> <sup>6</sup>	Positivist orientation – employing experimental design and statistical analysis in establishing causation	<ul style="list-style-type: none"> <li>- Rigorous means of assessing causal effects and seeking attribution with strong internal validity</li> <li>- Well suited to answer question of whether the programme works, thus being able to fulfil accountability purpose in evaluation</li> <li>- Avoiding selection bias</li> <li>- Providing the necessary assurance of certainty in a complex development context</li> </ul>	<ul style="list-style-type: none"> <li>- Limited in answering how and why the programme works</li> <li>- Limited in showing pathways of change and in explaining the change results for learning and improvement purpose</li> <li>- Weak in terms of demonstrating external validity</li> <li>- Ethical concern about RCT application in relation to people</li> </ul>	Results and experimental model <sup>7</sup>	<ul style="list-style-type: none"> <li>- Programme/project level</li> <li>- Ex ante evaluation design</li> </ul>
<b>Logical Framework Approach</b> <sup>8</sup>	<p>Considered to be a technocratic approach and operationalizes evaluation through two major elements: a chain of results with predefined objectives and corresponding performance indicators</p> <p>Dominant project planning, monitoring and evaluation tool with strong donor support</p> <p>Assumes a single truth that can be captured by an</p>	<ul style="list-style-type: none"> <li>- Structures and identifies programme logic and theory in the form of a results chain to show the expected causal links</li> <li>- Provides the technical parameters for various actors to agree on the assessment of outputs and outcome</li> <li>- Emphasizes upward accountability (with an audit focus) to the donor</li> </ul>	<ul style="list-style-type: none"> <li>- Criticised as overly simplifying and rigid</li> <li>- Its mainly quantitative performance indicators are limited in reflecting longer-term effects</li> <li>- Tends to neglect a broader range of unforeseen and unintended effects</li> <li>- Major focus on the role of the project implementation team in the actors chain, other actors tend to be abstracted</li> <li>- Thus limited in achieving learning given above</li> </ul>	Results, programme theory/theory-based model	<ul style="list-style-type: none"> <li>- Programme/project level</li> <li>- Mid-term and end-of-project reviews to measure output or outcome level results (better suited for short-term projects with limited, easily agreed and clearly defined results chain with simple and linear logic)</li> </ul>

<sup>6</sup> Adapted from: Bloom, 2006; Cartwright, 2007; Cartwright and Munro, 2010; Duflo et al., 2006; Jonas et al., 2009; Ravallion, 2009; Scriven, 2008; Shadish et al., 2008; Stern et al., 2012; White, 2013; White and Raitzer, 2017.

<sup>7</sup> This combines typologies from both Stern et al. (2012) and Hansen (2005).

<sup>8</sup> Adapted from: Armytage, 2011; Bakewell and Garbutt, 2005; Crawford et al., 2005; Earle, 2003; Gasper, 1997; Gasper, 2000; Hummelbrunner, 2010; Kothari, 2000; Lampis, 2005; World Bank, 2005.



	objectively verifiable indicator and not open to interpretation, thus a positivist approach		limitations		
<b>Outcome Mapping<sup>9</sup></b>	<p>Actor-centred evaluation approach</p> <p>Operates evaluation through its key concepts: outcomes defined as behavioural changes, boundary partners, outcome challenges, and progress markers</p> <p>The theory construction process and the identification of boundary partners may vary with different users and starting points, which makes OM a constructivist approach</p>	<ul style="list-style-type: none"> <li>- Recognizes clear system borders and the limit of project influence by claiming contribution instead of attribution</li> <li>- Enables reflection of process issues during evaluation through its evaluation mechanism and focus on behavioural changes of intermediary actors in the actors chain</li> <li>- Strives for a balance in achieving learning and upward accountability in evaluation</li> <li>- Enables internal accountability with its attention to organizational practices, and through fostering a sense of ownership and responsibility among staff</li> </ul>	<ul style="list-style-type: none"> <li>- Difficult to understand and apply its concepts and processes, especially when working in multiple languages</li> <li>- Demanding in time and resources to continuously engage project actors to fully achieve learning goal through regular reflections</li> <li>- Highly contextualized indicators and performance information gathered through progress markers may be difficult to aggregate</li> <li>- Main focus on intermediary actors and their behavioural changes may be inadequate and only reflects partly the picture of change along the results chain</li> </ul>	<p>Explanatory process, system, actor, programme theory/theory-based model</p>	<ul style="list-style-type: none"> <li>- Programme/project level</li> <li>- Best used from the project start, but can also be employed as an independent evaluation tool either during or at the end of the project for external or self-evaluation</li> <li>- Better suited to assess changes which require qualitative and contextualized indicators, in addition to those quantitative ones</li> </ul>
<b>Theory of Change<sup>10</sup></b>	<p>Theory-based evaluation approach, assessment through testing articulated assumptions and logical links</p> <p>Reliance on theory construction and interpretation as well as specific context makes ToC</p>	<ul style="list-style-type: none"> <li>- Provides a sound basis for assessment and adding explanatory value by making explicit how and why change might happen with the initiative – mapping out the logic and theory of underlying links and processes beneath a sequence of events (mechanisms)</li> </ul>	<ul style="list-style-type: none"> <li>- Its process is time and resource consuming</li> <li>- Possible confusion caused by the term ToC being used to mean very different things and by its diverse variations</li> </ul>	<p>Programme theory/theory-based model</p>	<ul style="list-style-type: none"> <li>- Programme/project level</li> <li>- Applicable in both long- and short-term projects</li> <li>- Flexible usage in mid-term or terminal reviews or assessment.</li> </ul>

<sup>9</sup> Adapted from: Earl et al., 2001; Hummelbrunner, 2010; IDRC, 2010; Roduner et al., 2008; Sheriff and Schuetz, 2010; Smith et al., 2012; Van Ongevalle, 2013; White and Philips, 2012.

<sup>10</sup> Adapted from: Barnett and Gregorowski, 2013; Centre for Theory of Change, 2016; Funnell and Rogers, 2011; INTRAC, 2012; Keystone, 2009; Stein and Valters, 2012; Vogel, 2012.

	a constructivist approach	<p>connecting outputs to outcomes or goal or impacts</p> <ul style="list-style-type: none"> <li>- Helps improve accountability (both upward and internal) and learning: through a reflective and adaptive learning system which fosters a mutual accountability and learning relationship among organizations and their counterparts</li> <li>- Enables deeper understanding of contextual factors</li> </ul>			
<b>Contribution Analysis</b> <sup>11</sup>	<p>Theory-based approach, using theory of change as main instrument in evaluation and seeking to demonstrate a plausible association (instead of attribution) between an intervention and observed changes and compile a substantiated contribution story</p> <p>Heavy reliance on well-informed interpretation and sense-making makes CA a constructivist approach</p>	<ul style="list-style-type: none"> <li>- Embeds system thinking and recognizes boundaries between various levels of results in the theory of change</li> <li>- Offers a pragmatic and systematic way to generate credible causal claims when other statistically strong methods are not feasible</li> <li>- Works well for understanding and interpreting changes in complex contexts involving multiple projects and actors</li> <li>- With its main evaluation instrument – theory of change – it shares similarities with ToC in terms of addressing accountability and learning</li> </ul>	<ul style="list-style-type: none"> <li>- Subject to the quality of evidence collected and interpretation of such evidence when compiling credible contribution stories</li> <li>- Necessitates strict quality criteria in applying CA procedures</li> <li>- Generated analysis may be challenged due to its heavy reliance on individual sense-making</li> </ul>	Programme theory/theory-based model	<ul style="list-style-type: none"> <li>- Programme/project level</li> <li>- Appropriate for complex interventions involving multiple projects, actors and various levels of results</li> </ul>

<sup>11</sup> Adapted from: Delahais and Toulemonde, 2012; Mayne, 2001; 2008; 2012a; 2012b; Patton, 2012; White and Philips, 2012.

<b>Political Economy Analysis</b> <sup>12</sup>	<p>Actor-centred approach</p> <p>Different variations can be explored for evaluation purpose (e.g. a rapid political economy assessment tool)</p> <p>Epistemologically rooted in rational theory</p>	<ul style="list-style-type: none"> <li>- Helps understand the political and economic context and the incentives of stakeholders, so as to identify feasible and realistic solutions to development challenges, and improve policy and programming</li> <li>- Helps explain and unpack the rationale behind the performance and achievement of the intervention during implementation process – this helps answer the question of ‘why changes happen’ – thus responding to learning purpose to some extent</li> <li>- Potentially helps serve benchmarking and (mainly upward) accountability function as it focuses more on the relationship and interaction between donors and country partners</li> </ul>	<ul style="list-style-type: none"> <li>- Narrowly focuses on an economic interpretation of politics or political factors and may neglect other prevailing contextual factors</li> <li>- Its strong tendency to reach technocratic solutions</li> <li>- The institutionalization of political thinking and working within aid organizations is challenging due to institutional constraints and cultural inertia</li> <li>- Findings and analysis from PEA remain to be better utilized for programme learning and improvement purpose</li> </ul>	Actor model	<ul style="list-style-type: none"> <li>- Multiple levels (e.g. programme/project, sector level)</li> <li>- Potential to serve as process evaluation tool alone (e.g. assessing political commitment and policy advancement) or in combination with outcome evaluation</li> </ul>
<b>Value for Money</b> <sup>13</sup>	<p>Assessment of 4 Es – economy, efficiency, effectiveness and equity – through economic analysis/appraisal (e.g. cost-benefit analysis and cost-effectiveness analysis) and results-based management tools</p> <p>Its major quantifiable and monetary means of measurement makes VfM a positivist approach to evaluation</p>	<ul style="list-style-type: none"> <li>- Helps provide different dimension to make informed project management decisions and judgement, in addition to statistical calculations</li> <li>- Adding VfM parameters to programme appraisal and impact evaluation has a significant effect on the design, implementation and evaluation of programmes</li> <li>- Enables implementing organizations to specifically demonstrate the value of their work and</li> </ul>	<ul style="list-style-type: none"> <li>- Overly dependent on individual subjective interpretation and application due to a lack of consensus on defining the approach</li> <li>- Risks such as short-termism in reviewing development results and possible exclusion of some more difficult contexts and projects with harder-to-measure aspects due to its focus on efficiency, costs and benefits,</li> </ul>	Economic model	<ul style="list-style-type: none"> <li>- Multiple levels (e.g. country programme, individual project level)</li> <li>- Both the means and the end product – a tool for evaluating and the object of evaluation (initial definition and clarification in specific case is necessary)</li> </ul>

<sup>12</sup> Adapted from: DFID, 2009; Fritz et al., 2014; Hudson and Leftwich, 2014; McLoughlin, 2014; Pettit and Acosta, 2014; Poole, 2011; Reich and Balarajan, 2012; World Bank, 2016; Yanguas and Hulme, 2014.

<sup>13</sup> Adapted from: Barnett et al., 2010; Bortcosh and Gibby, 2016; ICAI, 2011; Jackson, 2012.

		strengthen financial accountability to stakeholders	thus limited in realising learning as an evaluation tool		
<b>Cost-Benefit Analysis</b> <sup>14</sup>	<p>Regarded as a technocratic approach</p> <p>Possibly applied as a diagnostic and evaluation tool, and assesses through converting the project costs and benefits to comparable or quantifiable monetary units for systematic comparison to determine project worth</p> <p>Positivist approach with its quantifiable and monetary means of measurement</p>	<ul style="list-style-type: none"> <li>- Encourages people involved in policy development to consider cost and benefit aspects in making informed strategic decisions</li> <li>- Valuable in helping provide analysis to serve (financial) accountability purpose</li> </ul>	<ul style="list-style-type: none"> <li>- Its influence tends to be weakened by other factors such as political forces, conventional practices</li> <li>- Ethical concern arises when individuals' preferences largely influence decision-making process</li> <li>- Limited in accommodating and reflecting the non-commensurable dimensions of a policy or project in evaluation, therefore limited in achieving learning purpose alone</li> </ul>	Economic model	<ul style="list-style-type: none"> <li>- Programme/project level</li> <li>- Possibly used at different stages of the project cycle such as before, during and after project implementation</li> </ul>

These reviewed approaches in above Table 3.2 exhibit the diversity of options in the evaluation toolkit. For example, the identification of evaluation models illustrates different focus and questions to be pursued in evaluation by each approach. This table also shows how these approaches may respond to the key parameters set out in the conceptual framework. Some of them are more suitable for addressing upward/financial accountability in evaluation, such as CBA, LFA and VfM; some endeavour to balance accountability with learning, such as OM; ToC and OM both pay attention to intermediary mechanisms and links connecting outputs with outcomes or goal, and OM has specific elements enabling the assessment of such mechanisms or links; RCT is seen as a rigorous means of establishing causation in relation to the variable of interest; CA is specifically designed for the purpose of solving the issue of

<sup>14</sup> Adapted from: ADB, 2013; BetterEvaluation, 2014; Ergas, 2009; Hockley, 2014; Leong and Lim, 2016; OECD, 2006; Shapiro and Morrall, 2012; Webb, 1976.

attribution; and PEA has the potential to be further integrated into evaluation (e.g. process evaluation), and provide a valuable political dimension and measurement of accountability from an actor perspective. One evaluation tool can hardly fulfil all the parameters set out in the conceptual framework and provide an adequate solution to identified challenges and demands. Therefore, concerted efforts are needed to combine strengths and complement the limitations of different tools. One approach evaluators may often take is to mix and match different tools as necessary; another possibility is to synthesize elements of different tools into a single framework.

From this review and analysis, and among these selected approaches, the LFA and OM look promising with their various features, merits and demerits, especially in relation to the criteria in the conceptual framework: the LFA is a donor-supported dominant project tool and is well suited for satisfying upward accountability; OM takes an explicit actor perspective and adds certain explanatory power with its focus on intermediary actors and mechanisms between outputs and outcomes or goal. These two approaches merit a detailed review and close-up examination.

In practice, the aforementioned complementarity between the LFA and OM has been noted – there have been attempts in combining the use of the two in response to key evaluation challenges and a theoretical proposition about the synthesis of the two approaches has also been proposed (Roduner et al., 2008). Accordingly, in the following sections, I focus on and provide a critical review and analysis of the LFA and OM with reference to the constructed conceptual framework, and explore in depth their potential for synergy.

## 3.2 Critical review of the Logical Framework Approach and Outcome Mapping in evaluation

### 3.2.1 Critical review of the Logical Framework Approach in evaluation

#### 3.2.1.1 Overview of the LFA

The LFA is the most widely applied project management tool and functions through various stages from planning to monitoring and evaluation (Bakewell and Garbutt, 2005; Hummelbrunner, 2010, OECD, 2002).

The LFA was originally developed for the engineering and business management sectors and then introduced into the development field by USAID in the late 1960s (ECODE, 2011). It has established its status as the standard project management approach used by many multilateral and bilateral agencies such as the World Bank and DFID, and as the standard format required by many donors for grant applications (ibid.; Bakewell and Garbutt, 2005; World Bank, 2005).

Although there are some variations, the LFA is most commonly presented as a matrix containing its various elements, as shown below in Table 3.3.

**Table 3.3: The Logical Framework Matrix**

<b>Objectives</b> (Narrative summary)	<b>Objectively verifiable indicators</b>	<b>Means of verification</b>	<b>Assumptions</b>
<b>Goal</b> – overall aim to which the project is expected to contribute	Measures (direct or indirect) to show the project's contribution to the goal	Sources of information and methods used to show fulfilment of the goal	Important events, conditions or decisions beyond the project's control which are necessary for maintaining progress towards the goal
<b>Outcomes</b> (or objectives) – the new situation which the project aims to bring about	Measures (direct or indirect) to show what progress is being made towards reaching objectives	Sources of information and methods used to show progress against objectives	Important events, conditions or decisions beyond the project's control which are necessary if achieving the objectives is going to contribute towards the overall goal

<b>Outputs</b> – results which should be within the control of project management	Measures (direct or indirect) to show if project outputs are being delivered	Sources of information and methods used to show delivery of outputs	Important events, conditions or decisions beyond the project's control which are necessary if producing outputs is going to help achieve objectives
<b>Activities</b> – things which have to be done by the project to produce outputs	Measures (direct or indirect) to show if project outputs are being delivered	Sources of information and methods used to show that activities have been completed	Important events, conditions or decisions beyond the project's control which are necessary if completing activities will produce the required outputs
<b>Inputs</b>	Resources – type and level of resources needed for the project Finance – overall budget Time – Planned start and end dates		

(Source: Bakewell and Garbutt, 2005, p.3)

In this matrix, the first column depicts a hierarchy of levels of objectives and the causal logic among them; the second column identifies indicators of the fulfilment of objectives and targets related to each level of results in the first column; the third column shows the source of data related to performance indicators and how these data are collected; and the fourth column refers to those hypothesized conditions and influencing external factors which might affect the success of the project (Bakewell and Garbutt, 2005; Gasper, 2000).

The LFA is essentially a 'cause and effect model of project interventions to create desired impacts for the beneficiaries' (World Bank, 2005, p.13). Moreover, it is an analytical and presentational tool used to: facilitate situation analysis and identify various above-mentioned strategic elements (e.g. inputs, outputs, outcomes and impact; indicators; assumptions or risks) and their causal links; decide best strategies to monitor and evaluate changes such as outputs and outcomes; and provide a standard format to present a summary of activities if necessary (ibid.; AusAID, 2005, p.1; OECD, 2002).

The dominant status and wide application of the LFA notwithstanding, as briefly

mentioned in Section 3.1, it is better suited to 1) projects that are short-term and have simple linear logic with limited, easy-to-measure or observable goals and indicators; 2) projects that need to report upwards to donors for accountability purpose, and monitor and assess the implementation of planned project activities; 3) projects that focuses on a single organization's specific activities and results such as outputs and outcomes (Keystone, 2009, p.30).

However, donor organizations (especially those official aid agencies such as DFID) often impose the adoption of the tool to fulfil certain bureaucratic imperatives (Bakewell and Garbutt, 2005; DFID, 2011). As the Foundation for Advanced Studies on International Development (FASID, 2005, p.63) asserts, 'Most aid organizations'<sup>15</sup> have traditionally interpreted a project as a Logical Framework (LF) and evaluated it based on the prescribed indicators' and OECD evaluation criteria. NGOs or other implementing organizations working with these aid agencies are often required to apply the LFA throughout a project cycle irrespective of project type (Bakewell and Garbutt, 2005; Sida, 2004).

Despite strong support from the donor side and its prevalent usage, the LFA has drawn considerable criticism from both development researchers and practitioners. In addition to some general critical analysis in Section 3.1, the following sections further elaborate on those key aspects of the LFA in evaluation, with reference to the parameters set out in the conceptual framework.

---

<sup>15</sup> Such as the UN system, the German Organization for Technical Cooperation (GTZ), the Canadian International Development Agency (CIDA), USAID, the Norwegian Agency for Development Cooperation (NORAD), the Swedish International Development Cooperation Agency (Sida), the Australian Agency for International Development (AusAID), the European Commission (EC), and Japan International Cooperation Agency (JICA), and etc (Landoni and Corti, 2011; Sida, 2004, p.5).



### 3.2.1.2 Accountability and learning in evaluation with the LFA

According to Crawford et al. (2004, p.178), if an aid project is conceived as a social experiment, a LFA-based evaluation design tests three hypotheses which actually outline three broad mandates of development evaluation, as listed below in Table 3.4.

**Table 3.4: Hypothesis and forms of evaluation in LFA-based project and evaluation design**

<b>Hypothesis</b>	<b>Form of evaluation</b>
<b>Management hypothesis:</b> resources or inputs invested in the project by the donor will be efficiently converted to planned outputs by the project implementation team	<b>Management evaluation:</b> assessment of the efficiency of the project team
<b>Intervention hypothesis:</b> outputs actually delivered will foster the desired effects in the beneficiary community	<b>Intervention evaluation:</b> assessment of the efficacy of the project design
<b>Development hypothesis:</b> effects realized will contribute to the desired development impact	<b>Development impact evaluation:</b> assessment of the effectiveness of the development strategy or policy

(Source: Crawford et al., 2004, p.178)

Thus, the efficiency (outputs), efficacy (outcomes) and effectiveness (impact) of interventions can be assessed by the LFA through testing the respective underlying hypothesis. However, Crawford et al. (ibid.) raise a caveat to such a linear ‘chain of causality’: the criteria of all three hypotheses need to be met if the development impact defined in the LFA is to be achieved, which is somewhat challenging in the complex and emergent development environment. Moreover, this causal model implies that attribution can be feasibly established at various levels of change according to the linear logic of the LFA, which is seldom the case in reality.

The LFA operationalizes evaluation through its two major elements: predefined objectives (e.g. outputs, outcomes and impacts) and corresponding performance indicators. In evaluation, the LFA emphasizes upward accountability (with an audit focus) to the donor to demonstrate whether the intervention is delivering the results as

proposed (Bakewell and Garbutt, 2005; Gasper, 1997). Kothari (2000) also recognizes LFA's value in realizing the goal of (mainly upward) accountability in evaluation as it can structure a project well and establish the processes of accountability. The LFA is considered to be less suited for fulfilling internal and downward accountability purposes due to its imbalanced actor and result focus – to be explained in the following section.

However, the emphasis of upward accountability in evaluation with the LFA may create barriers to achieving learning. As Gasper (2000, p.26) argues, logframes 'represent a style of planning and evaluation that assumes high authority plus high levels of foresight; accountability considerations then predominate in ex post evaluation above learning considerations'. More directly, he points out that the audit form of accountability using the LFA can be at the expense of learning if major attention is focused on measurable indicators only (ibid.). Thus, learning and accountability may become even more dichotomous in evaluation with the LFA.

Furthermore, due to its rigidity and its instruments of evaluation against set objectives with predetermined and quantifiable indicators, the LFA has been criticized for leaving little room and flexibility for learning and adaptive adjustment in response to contextual influences, easily neglecting a broader range of unforeseen and unintended effects (Earle, 2003; Gasper, 2000; Harley, 2005; Hummelbrunner, 2010; Kothari, 2000). This limitation of the LFA is reinforced by below research findings of Bakewell and Garbutt (2005):

Where the LFA is used for monitoring and evaluation...the focus is often the logical framework – to look at the expected achievements laid out in the matrix – rather than the work itself... [Therefore,] the rhetoric of flexibility and learning which is suggested by the theoretical application of LFA does not work out in

practice. (pp.10–11)

Other authors such as Grove and Zwi (2008) note that the logframe matrix mainly reflects easy-to-define and measurable issues rather than the most critical ones, it thus tends to focus on the most tangible results.

Consequently, with these aforementioned problems of the LFA, there remain considerable barriers to achieving learning and to reconciling the tension between the two key goals when using the tool for evaluation.

### ***3.2.1.3 Actor consideration with the LFA***

Another problem identified with the LFA lies in its limitation in effectively involving a wide range of actors such as local actors and beneficiaries in a participatory manner (Chambers et al., 2009; Dale, 2003; Earle, 2003; Hummelbrunner, 2010; Kothari, 2000; Van Ongevalle and Fonteneau, 2014).

The rationale for this limitation is twofold. First, as the LFA is often imposed externally or from above and applied in a top-down manner, the power imbalance embedded in this process does not foster high trust, partnership or effective participation among project actors (Bakewell and Garbutt, 2005; Gasper, 2000; Earle, 2003; Hummelbrunner, 2010). Therefore, the views of local actors or lower-level staff are seldom collected during the evaluation process (Gasper, 2000).

Second, as briefly discussed in Chapter 2, the LFA tends to abstract project actors, focusing instead on the role of the implementation team in promoting various elements of the overall planned change (Crawford et al., 2005, p.6). The imbalance embedded in the way the LFA treats actors – heavy emphasis on the implementation

team and a neglect of other actors – may make evaluation with the tool ‘reinforced as a process to satisfy the donor requirements rather than a means of learning from the programme in a participatory way’ (Bakewell and Garbutt, 2005, p.11). This also helps explain that the LFA better fulfils upward than internal and downward accountability requirements. This imbalanced actor focus in the LFA also has implications for the explanatory power of the tool, to be discussed in the next section.

#### ***3.2.1.4 Explanatory power of the LFA – unpacking mechanisms linking outputs and outcomes or goal in evaluation***

It is recognized that the LFA is used more often and performs better in evaluating lower level objectives such as outputs, as it helps ensure sufficient resources for implementing project activities and efficiency in achieving outputs (Earle, 2003; Kilby, 2004). Moreover, as Armytage (2011, p.269) argues, ‘[Because the LFA] usually only measures indicators of process and efficiency, [it] has allowed many actors at the operational level to focus on ticking the boxes on the ‘deliverables’ in the ‘logframe’ as the day-to-day means of assuming [rather than actually demonstrating] developmental effectiveness’. This echoes the earlier observation of Kothari (2000, p.8) that the LFA does not seem to ‘offer real openings for the assessment of project outcomes as opposed to outputs’. On the other hand, it implies that the LFA may be good at demonstrating attribution at the output level in particular (which is arguably relatively feasible and straightforward, as mentioned in Chapter 2), rather than at a higher level in the results chain.

This problem gets worse with the imbalanced actor focus in the LFA (as discussed in the previous section) – leading to more attention being paid to outputs of project activities undertaken by the implementation team than other levels of results related to

different actors, as briefly mentioned in Chapter 2. This feature of LFA assessment weakens its explanatory power in evaluation as it may potentially cause a gap or missing links between outputs and longer-term results such as outcomes or goal.

In addition, as mentioned in Section 3.2.1.2, the LFA is also criticized for easily neglecting a broader range of unforeseen and unintended effects beyond its framework. Even in circumstances when such effects are considered, they are often not integrated into the original intervention logic (Hummelbrunner, 2010), let alone making clear the links between them (Couillard et al., 2009). This results in a lack of explanatory power in evaluation as the picture of change captured by the LFA may be incomplete and some important mechanisms and links between various levels of results may remain hidden and unpacked. This explanatory power is important if learning is to be fully achieved from evaluation by the LFA.

### ***3.2.1.5 LFA in practice – observations from selected examples***

Following the critical review and analysis of the LFA as an evaluation tool, this section draws on specific evaluation cases I gathered from the grey or organizational literature<sup>16</sup> to illustrate its application in evaluation and its responses to key parameters in the conceptual framework in practice. The following Table 3.5 summarizes each case in terms of evaluation object (project), funder, type, methodology, LFA use, and observations on how the LFA application may address the tension between learning and accountability, its explanatory power and its actor component.

---

<sup>16</sup> The availability of comprehensive, publicly available and high-quality documentation was one vital criterion for gathering these cases. Diversity of funding source, project type/sector, and project stage of evaluation were also taken into consideration in my case selection. These cases help exemplify LFA usage in practice, but there is a limitation of scope in terms of representativeness.

**Table 3.5: Case studies of evaluation with the LFA**

<b>Project evaluation case</b>	<b>Case study 1</b>	<b>Case study 2</b>	<b>Case study 3</b>	<b>Case study 4</b>	<b>Case study 5</b>
<b>Evaluation object</b>	In-service, Secondary School Teacher Education Project, Uganda	General budget support – the case of Burkina Faso (one of 7 country-level evaluations)	Community Based Psychosocial Support in Humanitarian Assistance, Sweden	Zikhulise Educator empowerment and curriculum materials development project	Pacific Horticultural and Agricultural Market Access (PHAMA) programme
<b>Project funder</b>	DFID	Multiple donors and partner governments	Sida	USAID	AusAID
<b>Evaluation type</b>	Independent, end-of-project evaluation (1999)	Joint (end-of-project) evaluation commissioned by donors <sup>17</sup> (2006), aiming for learning and addressing joint donor accountability at the country level	Independent and interim evaluation (2012), for accountability purpose and focusing on assessing outputs and outcomes	Summative evaluation (2001)	Independent review (2013) (end-of-project), to recommend whether the project should be extended to phase 2 and, if so, how it can be strengthened
<b>Evaluation methodology</b>	Reviewing or checking project outcomes mainly against pre-defined objectives and performance indicators in the logframe	<ul style="list-style-type: none"> <li>• Applying OECD evaluation criteria</li> <li>• Using possible sequence of effects laid out in the logframe as a basis for systematic test</li> <li>• Other evaluation data collection methods</li> </ul>	<ul style="list-style-type: none"> <li>• Combining theory of change and the LFA</li> <li>• Applying OECD evaluation criteria</li> <li>• Multiple data collection methods</li> </ul>	Using the LFA logic for evaluation but based on post-test only, evaluation against the LFA objectives and indicators	<ul style="list-style-type: none"> <li>• Following four evaluation criteria: relevance, effectiveness, efficiency and sustainability (largely following OECD definitions)</li> <li>• Impact assessment not included</li> <li>• Various evaluation data collection methods</li> </ul>
<b>LFA use in evaluation</b>	Using various levels of objectives in the project logframe, including outputs, purpose and goal, as well as objectively verifiable indicators as parameters for evaluation	A causality map indicating main causal links to be examined by the evaluation was developed on the basis of the sequence of effects set out in the logframe	<ul style="list-style-type: none"> <li>• Objectives and targets framed in the original logframe were used as a baseline for evaluation and reporting</li> <li>• Assessment against outputs and outcomes set out in the logframe, complemented by interviews with key staff</li> <li>• Using indicators in the logframe to guide the collection of outcome and impact</li> </ul>	Formal aims or objectives and quantitative indicators in the LFA were referred to during evaluation	Assessing progress towards the achievement of programme results as articulated in the logframe

<sup>17</sup> Donors commissioning the evaluation include: the governments of Australia, Austria, Belgium, Canada, Denmark, France, Germany, Ireland, Japan, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, the UK and the USA; the EC; the Japan Bank for International Cooperation (JBIC) and the Inter-American Development Bank (IADB); the International Monetary Fund (IMF); OECD/DAC; and the World Bank (Lanser et al., 2006).

			level data		
<b>Key observations related to the LFA with reference to the conceptual framework</b>	<ul style="list-style-type: none"> <li>Some outputs were not really achieved and some assumptions in the logframe did not seem valid</li> <li>Problem with articulation and suitability of performance indicators: specific and measurable features of indicators were not suited to assess learning element</li> <li>Upward accountability seems to be prioritized, and given above two problems, neither accountability nor learning purpose was fully achieved</li> <li>Problem of excluding both unexpected and unintended results or by-products beyond and outside the logframe</li> <li>This leads to limited learning and weakens explanatory power with incomplete results gathered and missing links among some results</li> <li>Implicit actor component</li> </ul>	<p>This potentially risks missing outcomes or results which were not articulated and captured in the logframe or in the causality map based on logframe. External influencing factors/assumptions were not elaborated in the causality map</p> <p>✚</p> <p>Tension between accountability and learning thus did not get reconciled due to the flaw in the logframe;</p> <p>✚</p> <p>Weak explanatory power due to possible missing results and links among them</p> <p>✚</p> <p>Implicit actor focus</p>	<ul style="list-style-type: none"> <li>the logframe lacked internal logic with weak linkage between outputs, outcome and impact.</li> <li>No indicator was set to track how improved capacity led to improved impact at beneficiary level.</li> <li>The logframe was found to be weak in defining how monitoring and evaluation can show the link between improved outcomes and impact</li> <li>Some clearer distinction should be made between theory of change and the LFA, and their respective roles in evaluation</li> </ul> <p>✚</p> <p>Above problems well suggest weak explanatory power of the LFA in this case and its inadequacy in addressing accountability (although upward accountability was prioritized) and learning</p> <p>✚</p> <p>Implicit actor component</p>	<ul style="list-style-type: none"> <li>Some innovative and important legacies/results of the project were identified only as unintended consequences or were overlooked due to a flaw in the LFA design</li> <li>Quantitative indicators were limited in reflecting the quality and effectiveness of the project</li> <li>The application of the LFA was flawed from the project planning stage, which resulted in problems in evaluation using the LFA elements</li> </ul> <p>✚</p> <p>Flawed design with the LFA results in ineffective response to accountability and learning requirements, and weak explanatory power</p> <p>✚</p> <p>Implicit actor component</p>	<ul style="list-style-type: none"> <li>Weak and incomprehensive logical model with unclear relationship or links between activities and results</li> <li>Performance indicators did not indicate progression from the baseline through milestones to higher level targets and key results</li> <li>Lack of performance indicators to reflect some effects (e.g. effects of increased export of primary products related to women)</li> <li>It was not explicitly indicated in the evaluation methodology framework that the LFA was used as one of the evaluation tools, but it can be seen that it served as the parameter with its expected results and indicators. The LFA itself was assessed as well, and problems were identified in this evaluation report</li> <li>✚</li> <li>Inadequate in addressing learning and accountability due to weak articulation of key elements and design of indicators; weak explanatory power with unclear links as mentioned above</li> <li>✚</li> <li>Implicit actor component</li> </ul>

(Adapted from AusAID, 2013; Enfield and Forsberg, 2012; Harley, 2005; Lanser et al., 2006)

It can be seen from these practical examples that several major aid agencies as well as some partner governments adopted the LFA in evaluation in different ways. All five cases utilized or referred to expected results defined in the logframe as either a baseline for assessment or a reference for building a causality map to guide evaluation. All cases other than Case 2 used performance indicators in the logframe as parameters to measure project progress or achievements.

In terms of addressing the tension between accountability and learning, in some of the cases, learning was given equal weight to accountability with the aim of generating lessons for improvement and informing the decision on project extension (e.g. cases 2 and 5). However, the LFA did not really help fulfil both evaluation purposes adequately due to identified problems with the validity and accuracy of performance indicators, and possibly missing and excluded results and links. Consequently, attribution and causal claims generated on such basis may be misleading and lack explanatory power. This observation corresponds to the literature reviewed in earlier sections. All cases indicate that there is an implicit actor focus in the LFA or in the way it gets applied.

From this review, it can be drawn that the LFA indeed helps establish the structure and process needed to achieve upward accountability in evaluation (although there were flaw in the logframe design in some cases). It can also be concluded that the LFA provides limited room for learning in practice – either not prioritized or not adequately achieved. It is also weak in term of explanatory power (can be seen from most cases) with gap or missing links caused by some uncovered or neglected results. In addition, the LFA needs to incorporate an actor perspective more explicitly to help resolve some aforementioned limitations. In general, the arguments and observations



gathered from both the literature and practical cases reviewed suggest that the LFA does not appear to respond well to key evaluation challenges and demands, as featured in the conceptual framework.

The limitations of the LFA notwithstanding, it remains dominant with strong donor support, this may be another factor causing unresolved evaluation challenges. However, it would not be practical to simply reject the approach, as ‘an outright rejection of the tool (logframe) would be highly displeasing for large donor agencies’ (Earle, 2003, p.8). This argument remains valid. It would be therefore an expedient alternative to combine or integrate the LFA with other approaches to address its inadequacies and utilize the strengths of different tools, as already mentioned in Section 3.1.

### **3.2.2 Critical review of Outcome Mapping in evaluation**

#### ***3.2.2.1 Overview of OM***

OM is a participatory project planning, monitoring and evaluation (PME) methodology developed by the International Development Research Centre (IDRC) in 2001 to respond to the weaknesses of existing approaches and challenges in monitoring, evaluating and reporting development results (Earl et al., 2001). OM introduces three innovative concepts: ‘sphere of influence’, ‘boundary partners’ – key actors with whom the intervention targets and interacts directly, and anticipates contribution to influence and change’ – and ‘outcomes’ defined as changes in the behaviour of boundary partners (ibid.). Specifically, OM focuses on: contribution and logical links and relationships between initiatives and outcomes rather than control and attribution; the complexity of the development context and process through recognition of the limit of its sphere of influence (ibid.; Jones and Hearn, 2009). From

this starting point, OM seeks to ‘help users to learn from and report realistically on their achievements by tracking the connections between what they do and what happens’ (Smith et al., 2012, p.11).

OM consists of three stages – intentional design, outcome and performance monitoring, and evaluation planning – and twelve steps<sup>18</sup> (as shown below in Figure 3.1).

**Figure 3.1: Three stages of OM**



(Source: Earl et al., 2001, p.4)

The role of OM as an evaluation tool has drawn considerable interest from both practitioners and researchers. Key elements such as outcomes, boundary partners, outcome challenges<sup>19</sup> and progress markers<sup>20</sup> are the major evaluation instruments of OM. The third stage – evaluation planning – was least developed in the OM manual

<sup>18</sup> A full glossary of OM terms involved in each stage can be seen in Appendix 1.

<sup>19</sup> Expected behavioural changes of a specific group of boundary partners and their relations to others if the programme/project achieves its full potential to facilitate changes (Earl et al., 2001).

<sup>20</sup> A set of graduated indicators showing performance and progress of boundary partners towards the desired outcomes. They advance according to three different levels: 1) what one would *expect* to see as an early reaction/response to the programme/project's initial activities; 2) what one would *like* to see at a more advanced stage; and 3) what one would *love* to see as a transformative state of change if the programme/project were to have a significant influence (Earl et al., 2001).

(as indicated in Figure 3.1), the usage of OM in evaluation thus largely depends on individual interpretation and adaptation in specific contexts.

OM is better suited to assess changes which cannot be captured by quantitative indicators alone. For example, Earl et al. (2000, p.109) state that OM is particularly helpful for providing deeper insights through ‘a qualitative, contextualized account of the development process’. Other practitioners argue that OM is ideal for capacity building related projects (Jones and Hearn, 2009). OM also has the advantage of potential for flexible application in evaluation, as reviewed in Section 3.1. This affords the possibility of synthesis with other evaluation tools in a single project.

In addition to this general overview of OM and some critical analysis presented in Section 3.1, the following sections further elaborate on those key aspects of OM in relation to the parameters set out in the conceptual framework.

### ***3.2.2.2 Actor consideration with OM***

As aforementioned major focus and core concepts of OM indicate, it has an explicit actor orientation, which differs from the LFA. Further difference lies in that OM focuses on key intermediary actors (boundary partners) with whom the project interacts or works directly, while the LFA pays more attention to the project implementation team. Each tool focuses on one type of actors in the actors chain.

This actor focus of OM has several implications. It adds value to evaluation by explicitly bringing an actor perspective which helps identify boundaries between different types of actors and then differentiate their respective roles and responsibilities (Sheriff and Schuetz, 2010). It also recognizes different perspectives

of actors working and operating in various domains and systems in a programme (Hummelbrunner, 2010). However, there are challenges in fully pursuing these values. It is difficult to maintain continuous interaction with project actors in practice after the initial intentional design stage as considerable time and resources are needed to sustain relationships and communication (Smith et al., 2012; Van Ongevalle, 2013). There is also potential for unequal power relationships between actors involved in the process of applying OM (e.g. when developing progress markers) (Sheriff and Schuetz, 2010).

Another key controversy lies in the inadequacy of the main focus of OM on behavioural change in direct boundary partners (Roduner et al., 2008). This may lead to a focus on only one part of the delivery and results chain at a time – results/outcomes as behavioural changes of boundary partners – with the consequence that crucial changes beyond those of boundary partners may not be accommodated in the immediate OM framework (Shaxson and Clench, 2011). Consequently, it is ‘difficult to build a strategic overview of the delivery chain, and to deal with complex networks of actors’ (ibid., p.5).

### ***3.2.2.3 Accountability and learning in evaluation with OM***

As briefly discussed in Section 3.1, OM strives for a balance in achieving learning and (multiple) accountabilities in evaluation (IDRC, 2010; Roduner et al., 2008). As the previous section mentioned, OM enables the identification and participation of relevant actors and seeks to engage them in dialogues to clarify objectives and agree on concerted action (Sheriff and Schuetz, 2010). In this process, a space is created for fostering a sense of ownership and responsibility (ibid.) among these actors, which

helps enable internal accountability and promotes collective project achievement for upward accountability requirement.

However, due to highly contextualized indicators and qualitative performance information gathered through OM instruments, it is difficult to quantify and aggregate data to satisfy donor preference for reading simple reports and information needs for upward accountability purpose (Roduner et al., 2008; Van Ongevalle, 2013). In addition, OM's main focus on behavioural changes in direct boundary partners at intermediary level may suggest that it is less capable of fulfilling downward accountability needs in relation to beneficiaries. The LFA shares this limitation.

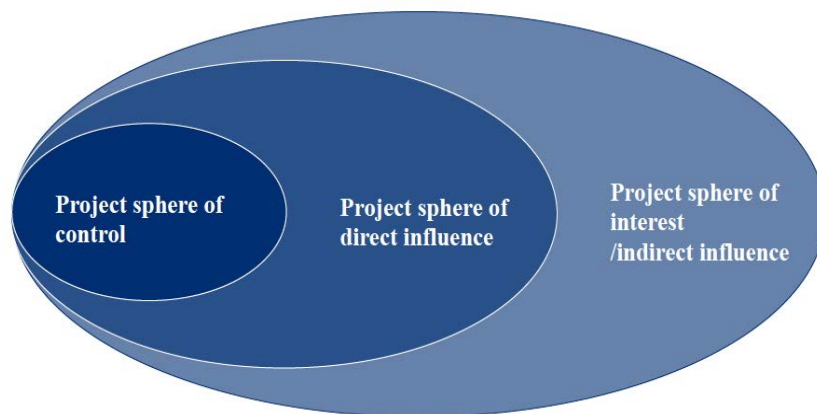
As Patton (cited in Earl et al., 2001, pp. viii-ix) comments in his foreword to the OM manual, 'The innovations introduced in Outcome Mapping provide ways of overcoming some of the barriers to learning faced by evaluators and development partners.' This is a distinctive advantage of OM in comparison to the LFA. As discussed in Chapter 2, an actor-centred perspective to evaluation helps enable learning through a more nuanced understanding of the characteristics and needs of different actors and tailored learning strategies. OM enjoys such advantage with a clear actor orientation. Furthermore, OM elements such as outcome, progress markers and strategy maps are useful instruments to unpack the process of change and learning pathways.

OM has the potential for reconciling the tension between accountability and learning in evaluation to some extent, but it needs continuous involvement and regular reflection among actors and necessary adjustment during its application. This remains rather challenging in practice, as mentioned in the previous section.

#### **3.2.2.4 Explanatory power of OM – unpacking mechanisms linking outputs and outcomes or goal in evaluation**

As mentioned previously, one of the key innovations of OM is its acknowledgment of the limit of a project's sphere of influence, as shown in Figure 3.2. OM focuses on a project's direct sphere of influence, but this does not mean it is limited to easier-to-achieve, less important or short-term goals. Rather, it actually 'focuses attention on incremental, often subtle changes, without which the large-scale, more prominent achievements in human well-being cannot be attained or sustained' (Earl et al., 2001, p.10). This feature of OM adds explanatory power to evaluation by looking at micro-level and intermediary mechanisms of change linking various levels of results.

**Figure 3.2: Sphere of project influence**



(Adapted from Earl et al., 2001; Hearn and Ambrose, 2013)

Specifically, as White and Phillips (2012) observe, OM tackles attribution through proving causal links (may be indirectly) between project activities and changes in behaviour, or in another words, connection between inputs/outputs and outcome gets unpacked. As outcome in OM is specifically related to actors who directly work with the project implementation team, it then can serve as intermediary mechanism/link connecting outputs with other levels of results (e.g. goal and impact) in relation to different groups of actors. By focusing and assessing outcome, OM meanwhile helps

unpack the mechanisms linking outputs with goal, and thus adds strong explanatory power to evaluation. In addition, there are corresponding elements of OM – progress markers – in place to specifically reflect those mechanisms (Sheriff and Schuetz, 2010) and a gradual and advancing process and pathway of change.

### **3.2.2.5 OM in practice**

Following the critical review of OM, this section draws on practical examples of OM usage in evaluation and relevant research study to exemplify its application and gather some observations on its utility, value addition, and challenges in its practical use by incorporating key elements in the conceptual framework.




There are relatively limited source of literature that directly focus on OM usage in evaluation. I tried to gather and examine some of those cases in the following Table 3.6 with the same criteria and procedures as those driving LFA case selection and analysis (see Section 3.2.1.5).

**Table 3.6: Case studies of evaluation with OM**

<b>Project evaluation case</b>	<b>Case study 1</b>	<b>Case study 2</b>	<b>Case study 3</b>	<b>Case study 4</b>	<b>Case study 5</b>
<b>Evaluation object</b>	Strategic framework of The International Land Coalition (ILC) (2011–15)	BioNET Global Programme (2007–10)	‘The Team’ Tanzania project	CARE Pathways programme	Six civil society projects in Bosnia and Herzegovina
<b>Project funder</b>	The International Land Coalition (ILC)	Swiss Agency for Development and Cooperation (SDC)	DFID	CARE, and Bill and Melinda Gates Foundation	Sida
<b>Evaluation type</b>	Independent mid-term review (2014), aiming for learning	Summative evaluation (2011), outcome rather than impact level evaluation	End-of-project evaluation (2013), aiming for both accountability and learning needs	Qualitative mid-term review (2015)	Annual review (2008), immediate outcome evaluation

<b>Evaluation methodology</b>	<p>Mixed-methods evaluation with OM as the major approach to capture outcomes shown by changes in behaviours and perception</p> <p>Focus on contribution rather than attribution</p> <p>The soundness of the original logframe was also reassessed</p>	<p>OM use based on evaluation standards of the American Evaluation Association</p> <p>Comparison of what was achieved through evaluation by OM with what was set in the logframe</p>	<p>Identification of outcomes in OM terms: behavioural changes of actors, evaluation data collected from methods such as case studies, document reviews, and informant interviews</p>	<p>Application of some OM concepts with other data collection methods such as focus group discussions and interviews</p>	<p>Prospective design, both baseline analysis and final assessment</p> <p>Evaluation methods included OM – comparison of data related to progress markers at baseline and final assessment stage – and in-depth qualitative interviews</p> <p>(This review also aimed to test OM potential as an alternative evaluation method for similar types of project)</p>
<b>OM use in evaluation</b>	<p>Major components of OM applied such as outcome definition, progress markers, and etc.</p>	<p>Adoption of definition of ‘outcome’ in OM to verify qualitative and quantitative progress towards objectives in the BioNET Business Plan and in the SDC logframe</p>	<p>Adoption of definition of ‘outcome’ in OM to track observable evidence that participants had actually applied learning or demonstrated attitude change in terms of changed behaviour</p>	<p>Key OM concepts such as outcome challenge and progress markers adopted to review the processes of change through focusing on behavioural changes</p>	<p>Adoption of OM to structure evaluation process with most key OM concepts such as outcome definition, boundary partners, and progress markers.</p>
<b>Key observations related to OM with reference to the conceptual framework</b>	<p>Application of OM to the whole process to measure behavioural changes would be difficult as ILC combines different actors with each making a distinct contribution and showing unique behaviour patterns</p> <p>Mapping changes of selected actors only does not reflect the complete change picture as changes among other actors will be missed</p> <p>OM can be seen as an option to deal with the complexity and unpredictability of results</p>	<p>This evaluation utilized OM notions to a large extent, such as focusing on results within the intervention’s sphere of influence; measuring outcomes (behavioural changes) rather than impact; and assigning contribution rather than attribution</p> <p>However, adherence to the OM outcome definition can be challenging</p>	<p>Some of the data did not show observable and verifiable behaviour changes (outcomes) but were descriptions of claimed changes in attitude, awareness, knowledge or capability</p> <p>More clarity is needed when using the OM definition of ‘outcome’, what is and is not included (e.g. observable and/or unobservable changes?), and how to collect data on unobservable changes in attitudes, relationships, etc.</p> <p>Bias mitigating measures should be strengthened</p>	<p>Use of OM concepts helps translate some abstract indicators into real-life, observable, and contextually meaningful changes, which further complements quantitative data obtained from other channels</p> <p>This shows that OM can be used in a flexible way.</p> <p>✚ This case focuses more on learning,</p>	<p>OM should be applied by donors from the project application stage, either using the whole package alone or together with the LFA</p> <p>Donors should focus on attribution more than contribution</p> <p>Some OM notions can be applied without formally implementing them such as thinking about actors involved</p> <p>This case provides a good example of a donor agency experimenting with OM in evaluation, and some valuable recommendations were given from donor perspective</p>



	 <p>This case shows efforts to prioritize learning in evaluation through applying OM elements with a clear actor focus, its explanatory power may be weakened as only some actors were selected and related changes reviewed.</p>	<p>This case shows an attempt to balance upward accountability (assessment against preset parameters in the logframe) and learning in evaluation by using key OM notions together with the LFA; OM use adds explanatory power by looking at behavioural changes of intermediary actors and their contribution to impact.</p>	 <p>This case aims for both accountability and learning, applying key OM elements to help examine specific learning process and outcome; stronger explanatory power through tracing evidences of change in relation to corresponding actors (in OM terms) in a detailed manner.</p>	<p>using key OM concepts to capture some qualitative performance information and mechanisms of change, thus adding explanatory power to other quantitative means of assessment.</p>	 <p>This case focuses more on learning purpose in evaluation and adopts an actor-centred perspective and other key OM concepts such as progress markers to reflect change process and mechanisms from baseline to final assessment stage, thus bringing explanatory power to the evaluation.</p>
--	--	--	--	---	---

(Adapted from CARE, 2017; Howard et al., 2011; Kishekya et al., 2013; Roebeling et al., 2014; Powell et al., 2008)

The above examples of OM use show that some major aid organizations have adopted or experimented with OM in either mid-term or end-of-project evaluations with the main aim of learning. All five cases focus on outcome evaluation with outcomes defined in OM terms, and most of them sought contribution rather than attribution. None of them applied the full OM package, preferring to choose some key elements – such as outcome, outcome challenge, and progress markers – together with other approaches. Generally, they suggest the value of OM in actor-oriented thinking, dealing with complexity, and its explanatory power through the tracing of micro-level change processes and mechanisms so as to achieve learning.

These cases also surface some challenges in applying OM, as revealed in the literature: it can be challenging for some project actors to comprehend and adopt OM definition of outcome; changes captured by OM from selected actors do not reflect a complete picture of change and risk missing changes among other actors and factors.

Additionally, the choice of OM for evaluation in some of my gathered cases was motivated by dissatisfaction with conventional assessment through the LFA. This does not mean the LFA was completely rejected, its instruments (e.g. preset parameters) were incorporated in some cases (cases 1 and 2). Furthermore, Sida (Case 5) explicitly recommended that donors should adopt OM right from the project application stage either as a comprehensive package or in combination with the LFA.

### 3.2.3 OM and the LFA

From the comprehensive review of the LFA and OM both in the literature and in practice, it can be seen that each approach has its place and is framed to serve a different evaluation purpose with varied scope – what each approach aspires to achieve and can achieve if implemented effectively. Table 3.7 below summarizes the evaluation role of the LFA and OM respectively, particularly in terms of the key parameters in the conceptual framework.

**Table 3.7: LFA and OM in evaluation**

	<b>LFA</b>	<b>OM</b>
<b>Actor perspective</b>	Implicit actor perspective; focusing mainly on the project implementation team and their activities and outputs	Explicit actor perspective; focusing mainly on intermediary actors whom the project implementation team works directly with
<b>Explanatory power (elucidating mechanisms linking outputs and outcomes or goal)</b>	There may be a missing middle link or hidden mechanisms between outputs and outcomes or goal due to imbalanced actor and result focus	Certain degree of explanatory power – instruments in place to unpack the micro-level mechanisms of change related to intermediary actors
<b>Accountability/learning</b>	Established structure, instruments and process for achieving upward accountability, but limited in enabling internal and downward accountability, and in achieving learning.	Striving for a balance between learning and multiple accountabilities but limited in fully fulfilling upward accountability, and inadequate for downward accountability.

(Source: The author)

From Table 3.7 and earlier discussions in this chapter, it can be drawn that OM may

be capable of resolving some of the key problems of the LFA through highlighting the qualitative dimension of change (Powell et al., 2008), and bringing an explicit actor perspective and certain explanatory power with a focus on intermediary actors and results. Nevertheless, it is unlikely for a single tool such as the LFA or OM alone to provide comprehensive solutions to the key issues reflected in the conceptual framework at once. A concerted effort is often needed to integrate the strengths of different approaches.

There have been constructive attempts to explore the possibility of combining the LFA and OM in practice from researchers, practitioners and donor agencies such as Sida and SDC as follows (Powell et al., 2008; SDC, 2011; Smith, et al., 2012):

- Using OM concepts in a LFA dominant project without explicitly naming them (OM definition of outcome should at least be adopted)
- Adding and identifying boundary partners in a LFA framework
- Using OM for internal planning, monitoring and evaluation, but reporting with the LFA to satisfy donors' requirements
- Synthesizing the LFA and OM into one framework and translating the language of OM into LFA terms (e.g. matching outcome in OM framework to one level of result in the LFA matrix)

These possibilities reflect two ways of combining the use of the LFA and OM, as briefly mentioned in Section 3.1: 1) a relatively flexible 'mix and match' approach which uses the two tools together; and 2) a synthesis which integrates elements of the two into a single framework or model. In terms of the second approach, there has been an earlier effort to integrate the results-oriented focus of LFA and process-oriented learning pathways of OM into one model (Roduner et al., 2008). In what follows I refer to this as the 'synthesis model'.

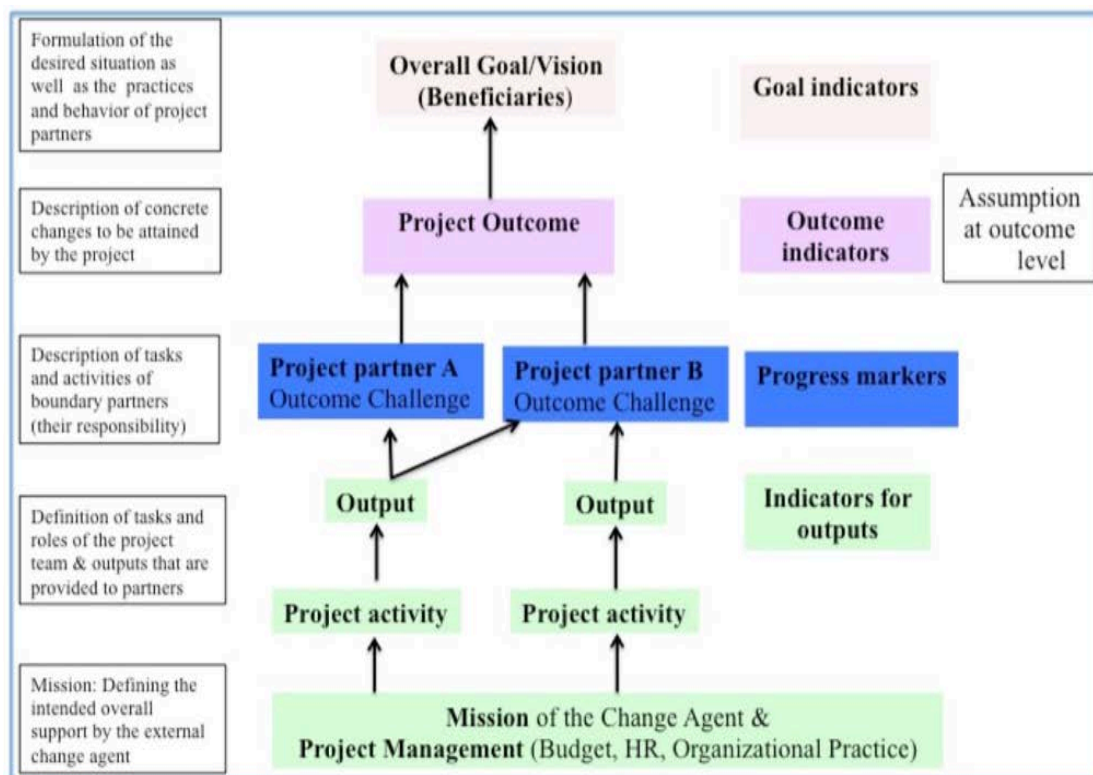
I distinguish the synthesis model from other ‘mix and match’ types of combined use of the LFA and OM in certain aspects. First, in terms of the underlying epistemological consideration, from a critical realist perspective, a synthesis can explore a middle way or strive for a balance between the positivist-oriented LFA and the constructivist-oriented OM. That is, it can reflect a relatively comprehensive picture of a stratified reality through evaluation. Second, from a methodological perspective, a synthesis approach may provide a shared space for fully integrating and absorbing valuable notions and the underlying logic of each approach in a more systematic manner. This corresponds to the typology observed by Stern et al. (2012), as reviewed in Section 3.1. Notions around evaluative thinking and design indicated in the conceptual framework – reconciling the learning and accountability dichotomy, actor perspective, and explanatory power with mechanism-based analysis – as well as results focus are expected to be fully embedded in a synthesis approach. The next section introduces this theoretically proposed synthesis model of the LFA and OM in detail, and explores its underlying theoretical basis, epistemological stance, and hypothesized potential.

### **3.3 The synthesis model of the Logical Framework Approach and Outcome Mapping: a potential evaluation alternative?**

Daniel Roduner and Walter Schläppi of the Swiss Association for the Development of Agriculture and Rural Areas (AGRIDEA), and Walter Egli of the Swiss Federal Institute of Technology (ETH) observed early on the potential for integrating the LFA and OM in a single framework to complement each other, and developed a synthesis model of the two approaches in 2008 (Roduner et al., 2008) (as shown in Figure 3.4 below). According to the authors, this synthesis model is designed to be the logic model that underlies a project from the planning stage onwards, and could possibly

take the form of a matrix similar to that of the LFA. It can be seen that there is an attempt to import key concepts of OM into this model, such as boundary partners, progress markers, changes in behaviour and capacities; and following the logic of OM, project outcomes are defined as results of behavioural changes in key project partners.

**Figure 3.4: Synthesis model of LFA and OM<sup>21</sup>**



(Source: Roduner et al., 2008, p.16)

With the conviction of the authors that the LFA and OM are not mutually exclusive but complementary, the rationale behind this design is to combine the advantages and strengths of both approaches and focus on both result and process. This synthesis model may be applied to different institutional contexts in a more comprehensive way. The designers leave the form of the model open and users can choose to

<sup>21</sup> Individual logical models of the LFA and OM can be seen in appendices 2 and 3.

emphasize either the use of LFA or OM elements according to specific contexts and the preferences of actors involved (Roduner et al., 2008).

The notions embedded in this synthesis model correspond to the views of Davies (2004) and Crawford et al. (2005) on the value of adding an actor dimension to the LFA. In an earlier work, Davies (2004, p.105) suggested that stages in logical frameworks ‘could be defined in terms of types of people as actors, in the sequence they were affected by and affecting others...[and] clearly defining the groups of people who are the actors in each stage of the Logical Framework can make the story line more evident and plausible’. This provides a sound grounding for the design of the synthesis model, which imports a clear actor focus from OM and links actors to different stages or levels of changes/results. Similarly, Dyer (2011) and SDC (2011) argue that OM concepts such as boundary partners, behavioural change and progress markers should be usefully assimilated into the LFA. The proposed synthesis model explicitly represents these proposals and provides a theoretical basis for testing them in practice.

Additionally, on the basis of LFA elements that represent relations between outputs and their effects, the model also seeks to include goal- or impact-level hypotheses and results and corresponding indicators to link project outcomes with the overall project vision and objectives (Roduner et al., 2008). This higher level of results – goals or impact – can be defined in conventional terms and reflect changes related to ultimate beneficiaries. The model therefore accommodates a more complete results chain compared with the single dimensional focus on behavioural change of OM alone. With outcome in OM terms included as intermediary level result, the mechanisms and links connecting outputs with goal can be then unpacked in evaluation to answer

questions of how and why changes happen, thus adding strong explanatory power. The proposed framework of the synthesis model helps reconcile donor preference for retaining the LFA to fulfil bureaucratic imperatives with a results focus and the need for explicit understanding of process issues during evaluation.

Drawing on the authors' proposed combination of the results-oriented focus of the LFA and the actor- and process-oriented learning pathways of OM (Ambrose and Roduner, 2009; Roduner et al., 2008), and the epistemological roots of both, I have the following considerations regarding the potential of the synthesis model:

- Firstly, I consider the synthesis model to be a theory-based evaluation tool that enables the development of a double-stranded theory of change – an actor strand and a results chain – for operationalizing either a comprehensive or a tailored evaluation<sup>22</sup>. In the case of complex programmes – multi-level entities as described by Cilliers (2001) – the model could also provide a framework for the breakdown of such an entity into various components, including ladders of results, relevant indicators and various groups of actors. The model therefore well accommodates the possibility of applying a mixed methods evaluation design<sup>23</sup>, as focusing on each individual component should require corresponding evaluation method rather than attempting to use one unified design for largely different components.
- Secondly, I consider that the synthesis model potentially reflects key characteristics of critical realism in evaluation, such as compatibility with a variety of methods, rejection of formulaic method prescription, and emphasis on mechanisms and their configuration with other influencing factors such as

---

<sup>22</sup> Focusing on and investigating the chain of programme theory in part or in its entirety distinguishes between tailored and comprehensive theory-based evaluations respectively (Weiss, 2000). The choice of a tailored or comprehensive theory-based evaluation depends on the specific evaluation focus, scope and purpose (Coryn et al., 2011).

<sup>23</sup> Mixed methods evaluations seek to 'integrate social science disciplines with predominantly quantitative...and predominantly qualitative...approaches to theory, data collection, data analysis and interpretation. The purpose is to strengthen the reliability of data, validity of the findings and recommendations, and to broaden and deepen our understanding of the processes through which programme outcomes and impacts are achieved, and how these are affected by the context within which the programme is implemented' (Bamberger, 2012, p.1).

context, as reviewed in Chapter 2. Furthermore, the synthesis model has the potential to help capture and reflect the stratified nature and the complexity of social reality in evaluation through a comprehensive system of components and mechanisms, including 1) various layers/levels of results focusing on different change dimensions (e.g. from micro behavioural changes to macro impacts); 2) its range of actors; 3) its contextual considerations; and 4) interaction and links between these components. This accords with the programme mechanism features discussed in Chapter 2.

- Thirdly, the synthesis model has potential for responding to those key parameters set out in the conceptual framework. As indicated by the design of the model, it seeks to offer a comprehensive package of indicators: quantitative (in the LFA style) and qualitative (e.g. in the form of gradual progress markers). There are also output-level indicators for verifying services provided by the project team (e.g. assessment of value for money). It thus has the potential for establishing an evidence chain and causal relationships through looking at minor behavioural changes in project partners (actor-centred) to explain how changes occur at a micro level and then contribute to macro-level changes over time; or in another words, potential for strengthening the explanatory power through unpacking the intermediary mechanisms connecting outputs and goal.
- Additionally, the model has the potential for reconciling the tension between learning and accountability: 1) learning to be potentially achieved through the process focus and added explanatory value of OM; 2) upward accountability to be addressed by the results orientation and evaluation instruments of the LFA, internal accountability potentially gets enabled with the identification and focus on those actors related to implementing organizations, downward accountability is potentially feasible as beneficiaries and related results are accommodated in the synthesis model but it needs considerable attention and efforts.

The synthesis model was proposed by its authors on a theoretical basis, and the potential that I highlighted in this section is mainly based on (ideal) hypotheses and assumptions. They need to be examined and verified through empirical testing. As



Patton (2011, p.7) comments, 'evaluation grew up in the projects'. It is therefore necessary to put this proposed model into practice and examine its applicability and practical value in evaluating development projects. This was basically my research starting point and the next chapter discusses how I put this theoretical model into practice with a detailed introduction to the research design and process.

## **Chapter 4 Research Design and Methodology**

First, this chapter outlines the overall research objective and major research questions. It then introduces the research design and methodology, including the overall case study strategy, evaluation design with the synthesis model, data collection process, and data analysis strategy and protocol. The chapter also includes a critical reflection on various key aspects in the research process.

### **4.1 Research objective and questions**

The overall rationale of this study is to further explore innovative, responsive and feasible alternatives for evaluating development projects under current evaluation context and development agenda. More specifically, my research objective is to pursue an alternative approach to evaluation through the practical application of a synthesis model of the Logical Framework Approach and Outcome Mapping with a case study approach. Accordingly, I expected that the applicability and value of the synthesis model could be examined empirically, useful lessons and implications drawn for future reference and research. I therefore formulated the major research question and sub-questions as follows:

What evaluative insights are afforded by integrating aspects of the LFA and OM to construct a synthesis model?

- 1) How does the synthesis model, as an alternative evaluation approach, respond to identified evaluation challenges and demands – in particular how well does it reconcile the tension between learning and accountability, offer explanatory power about the links between outputs and outcomes or goal and add an actor perspective?

- 2) What value is added and what limitations arise in adopting the synthesis model in practice?
- 3) What are the lessons and implications for further research and evaluation practice?

The following sections present the research methodology designed and adopted in practice to serve the research objective and answer the research questions.

## **4.2 Research design and strategy**

To fulfil the research objective, an appropriate research design is essential, and research content and objectives determine the methodology (Hart, 2005). Moreover, given the complex nature of development evaluation and its context, such research demands a collection of standard tasks (Pawson and Tilley, 1997). The desk review and secondary analysis in Chapter 3 help refine the research focus from a range of evaluation tools to the most salient of them – the LFA and OM – and then to a synthesis model of the two as a potential alternative evaluation approach.

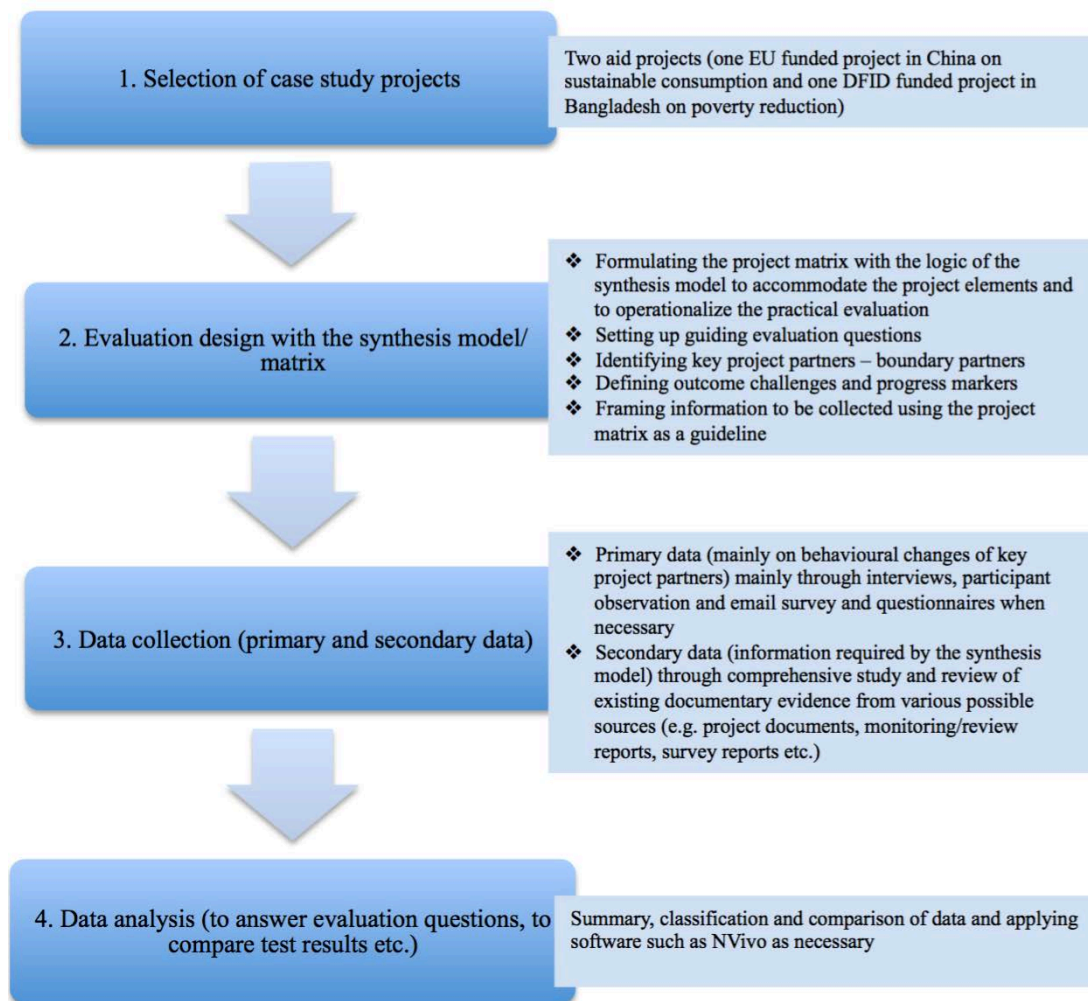
To test this theoretically proposed synthesis model in evaluation, and to examine whether the hypothesized potential of the model is realized in practice, evaluation objects – development projects – were needed to operationalize the research activities.

I thus adopted a case study strategy under which two development projects were selected for empirical research. The case study approach is defined as ‘an intensive study of a single unit or a small number of units (the cases), for the purpose of understanding a larger class of similar units (a population of cases)’ (Gerring, 2007, p.37). In spite of its recognized weaknesses in terms of representativeness and external validity, it was an appropriate research method for this study, as it enjoys ‘a natural advantage in research of an exploratory nature’ (ibid., p.39). It thus enables

intensive study and analysis of a small number of cases to generate some insights and lessons for future relevant research, especially considering the nature of the present study as well as the time and resource limits of a PhD project.

There were four stages in the research design, as shown below in Figure 4.1: selection of case study projects (prior to the fieldwork), evaluation design with the synthesis model and data collection (fieldwork research activities), and data analysis (post-fieldwork activity). During the fieldwork, I acted as an independent evaluator and conducted external, non-official evaluations of the two case study projects with the synthesis model. Due to the scope and capacity of a PhD research project, the synthesis model was adopted directly as a post-implementation evaluation tool at the project end rather than being introduced at earlier stages of the projects, as might be more desirable and practicable in a real-life deployment of the model. The following sections spell out the four stages in detail, including both my originally intended research process and the adjustments to it that proved necessary as I progressed.

**Figure 4.1: Overall research design and strategy (a case study approach)**



#### **4.2.1 A case study approach (rationale and selected projects)**

##### **4.2.1.1 Rationale for case study selection**

The selection of case study projects was critical in terms of operationalizing the testing and experiment with the proposed theoretical synthesis model as an evaluation tool. My selection was subject to many factors, such as availability of projects, willingness from the project side (both donors and implementation team), timing of approaching the projects, and potential appropriateness of project type. For example, the authors of the synthesis model assert that the model ‘is only applicable for projects in which capacity building plays a major role’ (Roduner et al., 2008, p.15), given its importance as an externally funded development intervention strategy.

However, it can be challenging to use the role of capacity building or development<sup>24</sup> as the criterion for defining the suitability of the synthesis model due to the lack of a clear and consistently recognized meaning of the term (Clarke and Taylor, 2008). I therefore decided it was more realistic to leave room for flexibility, which would also allow the possibility of examining the practical applicability of the model in different types of projects. Accordingly, I tried to ensure that I included one case which was most definitely a capacity building project by any definition, and one in which elements of capacity building were included but it was less central.

Additionally, as discussed in Chapter 3, behavioural change is a key element in the synthesis model as it adopts the OM definition of outcome as change in the behaviour of key project partners. Both behavioural change and the aforementioned capacity building are actor-centric elements and are closely connected to one another. As stated by Ambrose and Roduner (2009, p.2), ‘Changes in behaviour means strengthening the capacity of “local systems” (or their actors), which includes the capacity to continuously adapt and respond to a changing world.’ Therefore, I considered both elements as potential influencing factors when selecting case study projects, but they were not absolute criteria and there remained space for variation and innovation.

At the initial research design stage, I approached some projects and a few of them expressed interest and willingness to serve as case studies (including one UN project, two environmental international NGO-funded projects, and another two projects respectively funded by DFID and the EU which turned out to be the ones I selected as my case studies). My consideration and criteria for case selection can be summarized

---

<sup>24</sup> Capacity building and capacity development are often used interchangeably in the development literature and inter-governmental agreements and cooperation. The two are related, but capacity building is less comprehensive than capacity development (UNDP, 2008).

as follows:

- **Potential appropriateness:** 1) both aid projects were supported by major official aid agencies; 2) potentially suitable project types for the synthesis model as both involved behavioural and capacity change in key project partners.
- **Pragmatic considerations:** 1) these two projects were available as potential case studies at the right time, which was an important pragmatic consideration driving my selection; 2) appropriate timing for a series of empirical research activities as one project was finishing in the year of my fieldwork and the other had just finished; 3) other practical considerations – the first project was implemented in China and the project site was close to my residential area – which implied certain convenience in many aspects and it was time and resource saving relative to other options.
- **Potential value added:** Potentially widening the geographical coverage and adding more practical cases of OM usage, as OM has not yet been explored in-depth in China, although some research shows that its application has covered a wide range of regions (Smith, et al., 2012).

In addition to the above general considerations for case study selection, the following sections discuss the selected projects further in terms of specific background and context, existing evaluation system, donor evaluation requirements, and existing gaps and needs in evaluation that the synthesis model could possibly help address.

#### ***4.2.1.2 Case study projects: background and context***

The first case study (hereafter referred to as Case 1) was an EU-funded project aiming to promote resource efficiency and environmentally friendly economic development in China through mainstreaming individual sustainable consumption and improving the standard of living in the target area (two cities). With continuously improving living standards, Chinese citizens' consumption behaviour has changed and there is an increasing awareness of sustainable consumption (SC) and green supply chains.

However, although there is high degree of willingness to purchase green products, actual expenditure is much lower in practice. This is due to several factors, including availability, accessibility, price, and information about green products and related services (SWITCH-Asia, 2015a).

In such a context, Case 1 aimed to foster and mainstream sustainable consumption patterns through business networking among multiple stakeholders, thus creating a green product market by awareness raising and capacity building (e.g. for consumer associations) through an intensive EU–China partnership and experience exchange (BUCEA<sup>25</sup>, 2015). This project was piloted in two Chinese megacities with high population densities. It was expected that these actions could possibly be replicated and the sustainable lifestyle and consumption behaviour fostered in these two cities would play an essential guidance role for the rest of the country (ibid.).

As part of the EU's SWITCH-Asia grant programme – its latest regional cooperation strategy focusing on sustainable consumption and production – Case 1 generally followed EU requirements and guidelines in terms of implementation, monitoring and evaluation (BUCEA, 2012). The donor agency expected the evaluation to largely adopt OECD criteria and examine financial accountability (particularly important due to the grant element of the project); quality of project design; efficiency and effectiveness to date; impact prospects; potential sustainability; and cross-cutting issues such as gender dimension, environmental impact and governance (BUCEA, 2015). A learning element was also important in terms of effective dissemination of knowledge and information generated from the project for further replication purposes, but it was less explicitly stressed in the evaluation than accountability. In

---

<sup>25</sup> Beijing University of Civil Engineering and Architecture.



practice, a results-oriented monitoring (ROM)<sup>26</sup> system endorsed by the EU was applied as the main internal evaluation instrument and the original logframe was continuously revisited and adjusted on the basis of ROM results (BUCEA, 2012; 2013; 2015).

Available project information showed that the existing monitoring and evaluation data of this project were gathered mainly through a conventional results-oriented approach with the logframe as the main instrument (i.e. a linear activity–results model). There seemed to be potential for evaluation using the synthesis model to fill some intermediate gaps (e.g. between outputs and goal) through looking at the mechanisms and process of change from an actor-centred perspective. Thus, by selecting this project as one of the case studies in my empirical research, it not only made the testing of the synthesis model feasible, but also brought additional approach and perspective in reviewing the project results.

The second case study project (hereafter referred to as Case 2) was one component of a DFID-led programme in Bangladesh with the overall goal of achieving pro-poor economic growth in order to increase employment and income for the poor, especially women. Specifically, the project aimed to promote a sustainable micro-finance sector which could offer the ultra-poor, farmers, and micro and small enterprises greater access to and usage of diversified financial services (Khalily et al., 2013; PKS<sup>27</sup>, 2017).

---

<sup>26</sup> ROM provides an independent review at three levels: the micro level of the project, the macro level of EU development portfolio performance, and the meso level of the programme cycle. Its main focus is at the micro level, where it informs stakeholders of project performance and helps project managers think in results-oriented terms. It provides direct feedback on successes and problems during implementation and makes recommendations on how to improve operations (EuropeAid, 2012, p.22).

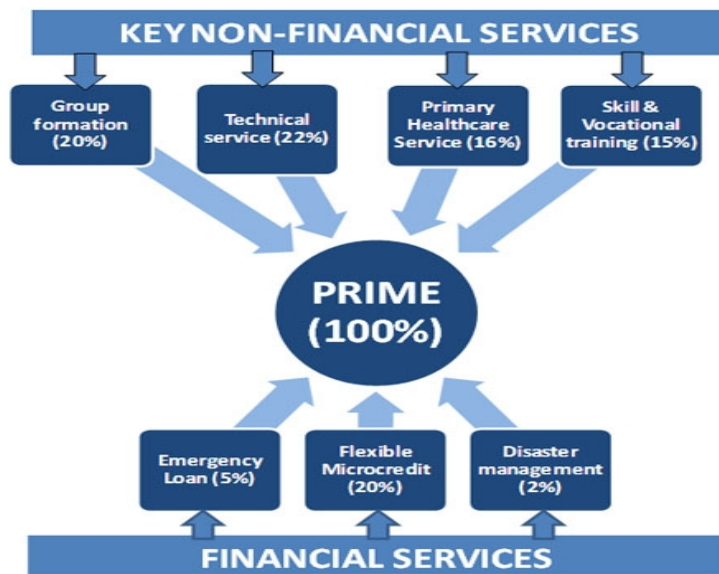
<sup>27</sup> Palli Karma-Sahayak Foundation, a Bangladesh government-supported development organization working on poverty reduction (PKSF, 2017).

Bangladesh, especially its north-western region, is stricken every year by a seasonal food shortage known as Monga<sup>28</sup> due to heavy reliance on agricultural production and the resulting vulnerability to climatic shocks and seasonal changes. During this period, people in the affected areas – the landless and wage labourers in particular – have minimal access to food and lack earning opportunities. In response, Programmed Initiatives for Monga Eradication (PRIME) – Case 2 – was introduced with support from DFID in 2006 (the first phase ended in 2013 and it was extended for a further two years) (Khalily et al., 2013).

The project aimed to prevent the negative consequences of Monga and Monga-like conditions, and relieve economic hardship amongst the ultra-poor in the short term and achieve poverty reduction in the long term. It primarily focused on creating employment and self-employment opportunities so that people affected by Monga could generate a sustainable income throughout the year (PKSF, 2013). As shown in Figure 4.2, PRIME had three financial components (flexible microcredit, emergency loan and disaster management) and four non-financial components (group formation, technical support, preventive primary healthcare, and skills and vocational training). PRIME was implemented across the northern, south-western and north-eastern areas of Bangladesh by a leading local development organization together with 24 frontline partner organizations (ibid.).

---

<sup>28</sup> A pre-harvest seasonal famine triggered by reduced availability of agriculture-related work (PKSF, 2013).

**Figure 4.2: Seven components of PRIME**

(Source: PKSf, 2013)

In its overall evaluation policy, DFID (2013) emphasizes that the key functions of development evaluation are generating evidence, learning what is working and what is not working, and promoting accountability. Case 2 therefore needed to follow these general guidelines in evaluation. It set up a systematic monitoring and evaluation framework, including a results-based monitoring (RBM)<sup>29</sup> system that reviewed results at output, outcome and impact levels (PKSF, 2013). Additionally, longitudinal impact studies and an issue-based short-term study were conducted (Khalily et al., 2013).

In spite of a relatively comprehensive monitoring and evaluation system in place in Case 2, there existed a missing 'middle': an instrument for assessing the role of the key capacity building element in the overall change process of the project, particularly

<sup>29</sup> RBM (as compared with conventional activity-based monitoring) was introduced to the project in 2010. The RBM system primarily focuses on results – output, outcome and impact – with the logical framework as its central element (PKSF, 2013).

in relation to its intermediary actors (24 partner organizations as mentioned earlier). Although I did not have access to official instructions from the donor agency specifically requiring an assessment of this element, one of the key respondents from the lead project team revealed such evaluation need during an interview. The adoption of the synthesis model in the evaluation of this project helped bridge the gap, particularly through the examination of outcome-level results, that is, behavioural and capacity changes in intermediary actors or boundary partners.

Following the overview of various aspects of the two case study projects, the following sections describe how I applied the synthesis model as an evaluation tool in terms of evaluation design, data collection, and the analysis process. The same procedure was followed in both cases throughout the entire research process.

#### **4.2.2 Evaluation design with the synthesis model**

This section introduces how I set out to apply the synthesis model to practical evaluation – reformulating the case study project frameworks according to the new model, and then using these frameworks as guidelines to identify evaluation data collection needs, channels and methods.

Specifically, with the research objective to explore the scope and applicability of the synthesis model as an evaluation tool, before judgements could be made and conclusions drawn about any value added, I first had to apply it to undertake evaluation in the two selected projects. I therefore needed to set up some evaluation questions: Did the project bring about any change or influence? What changes occurred and in whom? How did these changes happen? At the outset, my intention was first to apply the synthesis model to evaluate these two projects, and then

compare my findings and insights with those arising from official project evaluations that used the LFA, OM or other tools. However, my actual approach in reality had to diverge from this due to some practical constraints which I explain later in this chapter.

In order to apply the synthesis model as an evaluation tool to two case study projects that had originally been designed with other approaches, the first thing I had to do was to translate and reformulate original project data according to the synthesis model. At an operational level, to facilitate the usage of this theoretically proposed model and the translation of the original project information into its framework, I formulated a corresponding matrix with reference to the logic of the synthesis model, as shown below in Table 4.1, to accommodate essential project elements. In comparison to the original model in Figure 3.4 in Chapter 3, this operational matrix names and arranges outcome and goal level results in a slightly more nuanced way (as suggested by Ambrose and Roduner, 2009, p.3) to make it easier to apply in practice.

In the original synthesis model, outcome challenges are listed as intermediary results between project outputs and outcomes. However, this might be inconvenient during practical application for two reasons: 1) the terms ‘outcome challenge’ and ‘project outcome’ are used to refer to two different levels of result but the way they are named tends to conflate them; and 2) outcome challenges are specifically related to individual groups of boundary partners and if there is a large number of them, the presentation of corresponding outcome challenges might be difficult in a project matrix. Therefore, I grouped all outcome challenges collectively under the name ‘project outcome’ and used the term ‘project goal’ to replace the original ‘project outcome’ to make it easier to use and to differentiate key elements in the matrix more

clearly.

Additionally, as Table 4.1 shows, there are two elements listed as optional in the matrix, namely, the strategy map/journal<sup>30</sup> in relation to project activity, and project management and organizational practices. This is because these elements were included in the original design or logic of the synthesis model, but I did not explore them in my case studies because they entail much more in-depth involvement in a project from the planning stage, which was beyond the scope of this research project.

As previously mentioned, the key innovation of the synthesis model lies in the adoption of the OM definition of outcome as behavioural change in key project partners. It also adopts three other key concepts of OM: 1) boundary partners (those project actors with whom the project interacts directly and the project can expect to influence); 2) outcome challenges (how actors would behave and relate to others if the programme achieved its full potential); and 3) progress markers (a set of graduated indicators showing the progression of changed behaviour in a group of boundary partners) (Earl et al., 2001).

As neither case study project originally used OM, there were no actors pre-identified as boundary partners. Therefore, during the practical application of the model, my main initial efforts focused on identifying boundary partners, and setting up outcome challenges and progress markers during the translation and reformulation process. This necessitated a detailed preliminary document review, information extraction, and consultation with the project team to reach agreement on these key elements. This

---

<sup>30</sup> The element – strategy map/journal – is included in this project matrix with reference to both the original synthesis model by Roduner et al. (2008) and the updated version by Ambrose and Roduner (2009).

project framework reformulation process was conducted in each case study using the corresponding project matrix (the two completed project matrixes with most project elements are respectively shown in the case study analyses in chapters 5 and 6).

**Table 4.1: Reformulated project matrix of the synthesis model**

Project framework/matrix of the synthesis model			
Results/objectives	Content/statement		Indicators for measurement
<b>Overall goal/impact</b> (Beneficiaries)			Impact indicators
<b>Project goal</b>			Goal indicators
<b>Project outcome</b> (behavioural changes in key project partners/boundary partners)	Boundary partner A	Outcome challenge 1	Progress marker 1: Progress marker 2: Progress marker 3:
	Boundary partner B	Outcome challenge 2	Progress marker 1: Progress marker 2: Progress marker 3:
	Boundary partner C	Outcome challenge 3	Progress marker 1: Progress marker 2: Progress marker 3:
<b>Output</b>			Output indicators
<b>Project activity</b>			Strategy map/journal (optional)
<b>Project management and organizational practices</b>			Organizational practice (optional)

Following the reformulation of the project framework with the logic of the synthesis model, I used this new framework as an evaluation guide for identifying data collection needs and priorities so that an evidence chain could be constructed for final evaluation analysis. Evaluation data needs were identified in terms of four levels of results as listed in the above matrix: output, outcome, goal and impact.

After a preliminary screening and analysis of project information obtained at the initial stage, I determined different sources of data gathering in accordance with the

limited resources, time and capacity of a PhD research project: 1) primary data to be gathered on outcome-level results (in terms of actual behaviour change in identified boundary partners) using the outcome challenges and progress markers I had formulated as a guide; 2) data on output-, goal- and impact-level results to be gathered mainly through secondary data and information (e.g. from project documentation) with reference to various levels of indicators that I had formulated in the new project frameworks. In each case the interpretation I adopted for these levels of results in reformulating the project through the lens of the synthesis model, was as close as possible to that adopted in the original project logframe.

Finally, elements in the reformulated frameworks also informed my choice of evaluation data collection methods and the specific data collection design (this is further elaborated in the following data collection section).

#### **4.2.3 Data collection**

Given that one aspect of the assumed potential of the synthesis model is the possible accommodation of a comprehensive package of both quantitative and qualitative indicators, as well as a mixed methods evaluation design, this also means corresponding mixed (both qualitative and quantitative) methods of evaluation data collection during practical evaluation using the model. A variety of techniques, such as interviews, questionnaires, surveys, focus groups discussions, and participant observations can be employed to address specific evaluation questions (Bamberger et al., 2012; UNDP, 2009).

In practice, using the evaluation guidance set out in the previous section, I made preparations for collecting primary data on outcome-level results – behavioural



changes in key project actors – to address specific evaluation questions related to the change process. The semi-structured interview was the major method employed in my primary data collection. I designed interview guides with reference to the progress markers I had constructed in the project reformulation framework of the synthesis model, so that actual behavioural changes in key project actors (whom I identified as boundary partners) could be traced and relevant data collected (interview guides in relation to various types of respondents and an information sheet are shown in appendices 4–13).

Other than interviewing boundary partners identified in the two projects, I also interviewed other key actors including project team staff and donor agency representatives. A brief summary of interviewees in each case study in terms of numbers and roles is presented below in Table 4.2. I also conducted participant observation during project meetings and field visits to capture changing or changed behaviour and practice in the respective localities of the two projects.

**Table 4.2: A brief summary of interviewees in the two case studies**

Case 1			Case 2		
Type of respondent		Number of respondents	Type of respondent		Number of respondents
Boundary partner groups	Representatives from local consumer associations	7	Boundary partner groups	Project personnel from 22 partner organizations (aiming to include 1 management level staff member and 1 front-line project officer from each organization)	41
	Representatives from retailers	5			
	Representatives from suppliers	2			
	Representatives from local authorities	3			
Lead project team staff		7	Lead project team staff		13
Donor representatives		2	Donor representatives		1

Total number of respondents	26	Total number of respondents	55
-----------------------------	----	-----------------------------	----

When interviewing major respondents – key project boundary partners – I designed two sessions of questions for each interview. The first session included ten general questions mainly on overall project effectiveness from respondents’ perspectives, their observations as to the most significant changes, capacity and behavioural changes, before-and-after project situations, and so on. The second session covered a series of specific progress-marker-related questions aiming to trace the progress or achievement of project boundary partners (outcome) against the progress markers entered in the reformulated matrix (e.g. ‘expect to see’, ‘like to see’ and ‘love to see’). These progress marker questions essentially correspond to the rows in my reformulated project matrixes as presented respectively in Table 5.2 in Chapter 5 and Table 6.2 in Chapter 6.

I initially designed five specific progress-marker-related questions on actual progress; evidence and observation on what happened; influencing factors and challenges; the baseline situation related to a particular progress marker; and the follow-up plan for post-project activities. However, in practice, as this made interviews extremely long and interviewees’ time was limited, I was only able to ask two questions on actual progress, and influencing factors and challenges (the implications of this change in approach are reflected on in more detail in Section 4.3.1).

I basically followed similar structures in constructing specific interview questions in both case studies. The questions I designed for interviewing project staff and donor representatives in both cases were relatively simpler than those for boundary partners (as presented in appendices 9, 10, 12 and 13). In response to a suggestion from the

project team leader in Case 1, I adjusted my data collection strategy slightly by first using questionnaires to collect an initial round of data during a project summary meeting when seeing some key project partners for the first time. After establishing initial contact with them, I followed up with face to face or telephone interviews as they preferred.

My primary data consists of audio recordings of 3 recorded interviews for Case 1 and 52 for Case 2 (there are fewer recorded interviews for Case 1 as the progress-marker-related questions were relatively simpler and note-taking was therefore easier, especially in my native language). There are interview transcripts for all 26 respondents in Case 1 and 10 interviews transcribed from 55 interview notes for respondents in Case 2. I did not transcribe all the interview notes in Case 2 as there were many more interviews than in Case 1, it would have been very time-consuming to transcribe all the interviews, and my ultimate research purpose was not to compile an evaluation report based on the whole dataset (this strategy is discussed further in the reflection section of this chapter). For two groups of boundary partners in Case 1 – retailers and local consumer associations – I completed a data analysis matrix (see tables 5.3 and 5.4 in Chapter 5) into which I entered the assessments of each boundary partner against the corresponding progress markers.

In addition to primary data collection that mainly addressed the outcome level of the synthesis model, I sought additional information on other levels of results (outputs, goal and impact) through available secondary data sources. I also explored factors related to the contexts in which the projects were implemented – social, economic and institutional settings, and policy environment – through both secondary data (documentary reviews) and participant observation to provide critical information for

in-depth understanding and analysis.

Finally, in addition to data collected for evaluation purpose according to the synthesis model and matrix, I also gathered some informal views and comments from a few participants on their experiences of the evaluation process with the synthesis model. These are incorporated into the reflections on the research process in Section 4.3.

#### **4.2.4 Data analysis**

I then summarized, categorized, integrated and compared data (changes/results along the results chain of the synthesis model) collected by various methods from different sources to answer the evaluation questions, fulfil the evaluation objectives, and help draw conclusions on the test results of the model. More specifically, secondary data collected from documentary sources provided background knowledge and reference points, and the information required to complete the synthesis model framework/matrix; and was also important (with regard to both quantitative and qualitative indicators) in addressing evaluation questions which could not be answered with primary data.

I analysed primary data mainly gathered from interviews, questionnaires and observation to 1) help present the complex change process from the micro level (e.g. behavioural change in intermediary project actors) to the macro level (goal/impact at the beneficiary level) in order to explain how change occurred cumulatively over time; 2) illustrate how the capacity of those involved in project implementation had been gradually built with a view to ensuring sustainable change in beneficiaries; and 3) identify space for further learning and adaptation. I then integrated the findings from these analyses with other available secondary data related to project

performance and effects to formulate some evaluation results of the case study projects according to the framework of the model.

As mentioned earlier, according to my original research design, I intended to compare the findings of my evaluation using the synthesis model with those from official or other forms of formal evaluation in the two case study projects. However, this did not prove to be feasible. My empirical research was conducted at a stage in each project's lifecycle when results of official and formal terminal evaluations were not yet available. Additionally, there were also some restrictions from the project side in terms of sharing evaluation findings<sup>31</sup>. In these circumstances, I therefore sought to analyse the data and information available in project monitoring reports, interim review reports, and periodic impact assessment studies (available only for Case 2), to develop a sense of how my evaluation using the synthesis model related to what was available in the way of existing and emerging evaluation findings. Some examples of the results are presented in the following two chapters.

I also attempted to compare the evaluation mechanism of the synthesis model with that of other methods such as the LFA alone to illustrate its usefulness, strengths and weaknesses, together with limited interview participants' views and experiences with the model. However, such comparison was certainly subject to the evaluation-related documentation I could gather. I obtained some internal review reports in Case 1, and some monitoring and longitudinal impact study reports in Case 2, but did not have access to official and terminal evaluation results. A strict comparison was therefore not possible in practice.

---

<sup>31</sup> I made further attempts in this regard at the stage of writing up this thesis, but such restrictions in sharing information on official evaluation remain.

#### 4.2.5 Case study analysis protocol and justification of analysis strategy

This section discusses the specific strategy I adopted to present and analyse the two case studies during the application process of the synthesis model. For each project in turn, this protocol 1) presents a holistic project reformulation matrix using the logic of the synthesis model and introduces the reformulation/reconstruction process; and 2) follows a specific analysis strategy with a sequence of actor analysis, results chain analysis, and evaluation mechanism analysis<sup>32</sup> under the framework of the model.

The questions I sought to answer with this strategy were **who** are the major actors and focus in evaluation design (an actor-centred perspective), **what** specific results/changes are examined, and **how** are those changes evaluated (the instruments and mechanism deployed for guiding and operationalizing evaluation data collection and interpretation). Using this case study analysis strategy, I expected to be able to unpack the change process through the breakdown of key components in the model, as well as illustrate the evaluation role and utility of the model with examples from the case studies. This helps provide empirical basis for the final appraisal of the synthesis model (e.g. with reference to the key parameters set out in the conceptual framework) in the concluding chapter.

As discussed in Chapter 2 and reflected in the conceptual framework, an actor-centred perspective is valuable in providing insights into and understanding of the change process and effects through revealing various micro–macro relationships and interactions between external intervention, and local context and actors. By employing an actor-oriented approach as a major case study analysis strategy, I intended to make explicit the chain of actors involved at different stages of change as

---

<sup>32</sup> The term ‘evaluation mechanism analysis’ is used to refer to the analysis of the process and means involved and deployed in guiding and operationalizing evaluation data collection and interpretation.

well as their interactional relationships during the intervention, and then to unpack the pathway of change from minor behavioural changes in individual groups of project actors to macro-level results. In addition, ‘boundary partner’ is one of the key innovative concepts imported from OM to the synthesis model and it indicates the centrality of the actor dimension.

Furthermore, my underlying consideration of results chain analysis is not only to formulate a logical sequence of analysis which answers the question ‘what’ (changes are examined) following the identification of ‘who’ (the major actors are) as previously mentioned, but also to stress the weight and key status of the results chain in understanding the underlying programme theory and ladders of change which occur.

According to Funnell and Rogers (2011, p.177), the results or outcomes chain ‘shows the assumed or hypothesized cause-and-effect or contingency relationship between immediate and intermediate outcomes and ultimate outcomes or impacts (both short and long term)’ and is located ‘at the heart’ of programme theory<sup>33</sup>. It serves as the main instrument for articulating how the programme will work to achieve results and ‘it is (or should be) the centerpiece for developing all other aspects’ of the programme theory (ibid., p.179). Adopting a results chain in evaluation could help ‘avoid excessive focus on activities and getting stuck in activity traps that give insufficient attention to outcomes’ (ibid., p.180).

In adopting the results chain dimension as a device for analysis, I intended to make

---

<sup>33</sup> According to Funnell and Rogers (2011, p.13), programme theory is an explicit statement of how change will occur and how an intervention will produce causal processes; and it consists of theory of change and theory of action. The results or outcomes chain is a key component of theory of change in programme theory.

explicit the ladder of change/result in relation to various actors at different project stages and the relationships among them. This constituted another key step in uncovering and deconstructing the change process with a further breakdown of the results component. I then proceeded to analyse the mechanism for evaluating these changes and interpreting the evaluation data under the framework of the synthesis model.

### **4.3 Reflections on the research process**

This study is explorative in nature and there have been some challenges along the way from the initial research design (especially when seeking case study projects) to the data analysis stage. As mentioned in Section 4.2.1, I managed to gain access to two case study projects from major official donor agencies, which enabled me to explore empirically whether the synthesis model realizes its assumed potential in practice, and to fulfil the research objectives. In addition, timing to apply the synthesis model was also critical (although it would be best to apply the model from the project inception stage, it was beyond the capacity of my PhD study) – both projects were finishing at the time of the fieldwork. In Case 1, I was able to participate in the final project summary meeting, which most key project stakeholders attended, and this certainly provided a convenient opportunity to approach key respondents to arrange further interviews.

I basically stepped into each project as an external, independent researcher and evaluator who had no prior project-related role or involvement. This had both advantages and disadvantages. The advantages mainly lay in the fact that I could maintain a comparatively distant and independent status and perspective when interacting with project staff members, donor agencies and other actors. However, I



certainly had less in-depth understanding of the project rationale, inception process, or political dynamic due to a lack of deep involvement in the project.

The following sections reflect on the research process in a more critical and reflexive manner and look back at the limitations and challenges encountered during the research journey in terms of the evaluation design, and data collection and analysis process, other influencing aspects as well as ethical issues.

### **4.3.1 Reflection on limitations and challenges**

#### ***4.3.1.1 Methodological reflection***

- 1) *Evaluation design*: As mentioned earlier, since I applied the synthesis model as a post-implementation evaluation tool in a retrospective manner rather than from the project planning and implementation stages, it involved project reinterpretation and reconstruction (according to the synthesis model), boundary partner selection, and outcome challenge and progress marker formulation on the basis of information extraction, translation and inference from original project documents. During such a process, issues such as inaccuracy and bias may arise. For example, selecting boundary partners (who were also vital interviewees) was critical, and different evaluators, or different project actors, might have selected different boundary partners. The categorization of progress markers into ladders of change such as ‘expect to see, like to see and love to see’ can be a subjective process as well. To mitigate these potential issues, I sought to first gather as many original project documents as I could and refer to them as precisely as possible, and then to systematically triangulate by ensuring I interviewed a wide range of project personnel (as reflected in Table 4.2). I also consulted experts on OM in the field, and the OM manual for reference.

- 2) *Data collection process*: Various levels of identified progress makers served as the guideline for primary data collection through interviews with targeted respondents. As discussed in Section 4.2.3, I was unable to fully utilize my original comprehensive set of interview questions. In practice, the interviews proved lengthy and too detailed for most respondents – boundary partners – to answer all the questions. Practical restrictions such as language difficulties and time constraints also made it unrealistic to pursue the entire original interview guide with more respondents, although I did try to compile a full set of responses to my initial interviews conducted with just two individuals. I therefore redesigned and simplified the interview guide to ask just two questions about each progress marker, as mentioned earlier.

Furthermore, questions on the views of relevant respondents (both project team members and boundary partners) about the synthesis model were initially included in the interview guide. However, only a few respondents were actually asked these questions as it had already taken considerable time to explain the rationale of the interview, some basic information about the model, and the main interview questions. Another constraint was the fact that the model might appear complicated (especially OM related concepts) and thus difficult for most respondents to comprehend in short space of time. Therefore, only a few experienced project professionals were able to grasp the essence of the applied model and give limited comments in this regard.

When designing the interviews for primary data collection, I also considered an alternative survey element which would ask respondents to score and rank level of

progress or accomplishment made against set progress markers. I expected that I would later quantify the result on a numeric scale and augment it with a narrative description of changes, contributing/hindering factors and actors, sources of evidence, lessons learnt, and adjustments required. However, this element was not included in practice as the focus was put more on gathering data on actual changes in a qualitative format to supplement the quantitative indicators set by LFA element in the synthesis model. Such an approach would perhaps be more useful if seeking an easy way to aggregate evaluation data for a simpler report.

- 3) *Data availability, quality and coverage*: Applying the synthesis model as an external and ex post project evaluation tool as an ‘outsider’ made it somewhat challenging to gain access to core project documents and materials. This created some barriers to the accurate reformulation of project frameworks with the synthesis model, particularly in setting up core elements such as outcome challenges and progress markers. Additionally, as briefly mentioned earlier, due to time constraints and restrictions in the sharing of information related to evaluation, I had limited access to key official reports and findings in both case study projects. This inevitably limited the extent to which I could compare results generated from the synthesis model with those of formal project evaluations.

In addition, respondents tended to stay within their comfort zone in interpreting and answering interview questions. It therefore took time to engage them with the aim of the questions. For example, it happened quite often that when I asked about intermediary- and institutional-level changes (e.g. behavioural or capacity changes), respondents tended to answer in the first instance about changes at the beneficiary level. This revealed the need to shift their thinking from mainly the

‘demand-side’ to the ‘supply-side’ at a certain point. This situation resulted in some less relevant answers and data gathered.

Issues such as staff turnover and unwillingness to cooperate were also barriers to the gathering of quality and in-depth data. For example, a key staff member from one of the identified boundary partner groups left the organization at the time of the fieldwork, which led to the lack of an in-depth view and an information gap. Other staff members in this boundary partner group were reluctant to become interview respondents as they felt less confident in providing relevant information. To tackle these data issues, I traced and utilized some available monitoring data to supplement and verify evaluation data obtained from primary sources.

- 4) Evaluation data analysis: The data analysis stage was quite challenging, particularly in terms of effectively balancing my role and shifting it from evaluator to researcher. During the fieldwork, I approached relevant organizations and respondents as a research student but performed mainly as an independent evaluator when designing the evaluation with the synthesis model and collecting corresponding data. Having finished the fieldwork and gathered various data on project performance and effects, I needed to utilize and analyse those data to fulfil my research objectives.

By the time I started to analyse my data, I began to see the need to adjust my analysis strategy for several reasons. First, I recognized that as I was not the official evaluator of these two case study projects, the time it would take to complete full transcription of interview data into data analysis matrixes was not

justified because I would not be using them to write detailed evaluation reports.

Second, I knew by this point that I could not obtain official evaluation data (e.g. from independently commissioned terminal evaluations) for either case, so my initial aim of comparing the findings from official final evaluations with those of my own ‘simulation’ evaluation was no longer feasible. I therefore stopped inputting individual respondents’ data into the data analysis matrix. Using the secondary data I could gather from monitoring reports, mid-term evaluation reports, interim review reports, longitudinal impact studies, etc., I compared the data contained in them with the evaluation data I had generated through applying the synthesis model and especially those from my primary research. It was therefore a partial rather than a comprehensive and exhaustive exercise in terms of coverage. Nevertheless, I consider that it still provided a sound basis for me to make judgements on the applicability of the synthesis model in evaluating each of the case study projects, and over and above these two case study projects, to draw conclusions in respect of my research questions.

Finally, the data analysis process mainly involved interpretation and sense making from my standpoint as a researcher, which was inevitably subject to potential bias and certain limitations. I acknowledged this and mitigated the risk by using supplementary data from various sources (e.g. monitoring or evaluation related information from original project documents) to verify and triangulate my findings. It would have been more helpful if I had conducted more initial analysis at the time of data collection, which could have helped feed back to the data gathering process and relevant adjustments could have been made.

The above reflections are mainly on technical challenges I encountered during the

empirical research in the first level; later in the concluding chapter of this thesis, I apply and utilize some of the evaluation data to the wider, ‘meta-analysis’-type task of responding to my research questions and reflecting on a series of other challenges and limitations related to the synthesis model itself and its application.

#### ***4.3.1.2 Reflection on other aspects of the study***

Other challenges were related to logistics and the local context during the fieldwork in Bangladesh in particular. For example, one of the scheduled key informant interviews was cancelled due to a strike in Dhaka. Other influencing factors such as institutional setting and features were also significant. Given that local consumer associations (one of the identified boundary partner groups) in China function mainly as government agencies rather than civil society organizations, the bureaucratic mechanism needs to be taken into consideration when conducting research with them on subjects related to evaluation in particular.

Moreover, the preferences and openness of the funding agencies of the two case study projects also affected the data collection process to a certain extent. In this regard, DFID showed more interest than the other donor agency in learning the effects of capacity building in addition to other aspects of project performance and results. This was evinced by easier access to interviewees and some key project documents such as impact assessment reports.

#### **4.3.2 Reflections on ethics in the research process**

Ethical issues were attentively considered in the whole research process and ethical guidelines and principles set by the University of Sussex were followed. Before commencing my fieldwork, I fulfilled the formal ethical review procedures as

required by the University, including elements such as general fieldwork checklist and risk assessment, data collection and analysis plans, checklist on procedures in relation to informed consent and recruitment of participants, confidentiality and anonymity, fieldwork context, as well as relevant supporting documents<sup>34</sup>.

After getting formal ethical approval from the University, I specifically adopted the following ethical protocols and conduct as guideline throughout my field research process, especially when conducting formal interviews and questionnaires:

- informing respondents/informants my research area, purpose and agenda
- getting informed consent before starting interaction and collection of data
- getting permission (mostly verbal permission as preferred and agreed by interviewees) of all forms of recording and informing participants of their right to withdraw from information sharing at any point during or after the process
- informing respondents of the data storage and use (the data to be stored securely without being accessible to others)
- ensuring confidentiality and anonymity of respondents

These key principles were applicable to most of my fieldwork settings and methods for data collection. In circumstances when I conducted participant observation at project meetings or project sites, in order not to interrupt the process, I mainly pursued verbal consent from the participants on possible information usage later in my work.

As Patton (2002, p.407) notes, qualitative methods such as (in-depth) interview are highly personal and interpersonal – the researcher enters into the lifeworlds of people and tries to ‘open up what is inside people’ – such inquiry or intervention (Mikkelsen, 2005) can be therefore more intrusive. Related ethical issues need to be considered.

---

<sup>34</sup> These supporting documents contain some personal information. I therefore did not append it, in accordance with the General Data Protection Regulation.

Generally speaking, when I was conducting interviews during my fieldwork, I maintained a very careful awareness that I was taking up the work or life time of my respondents and that my research activities might be perceived as kind of ‘intrusion’ into their lifeworlds from respondents’ perspective. I tried to take all reasonable measures and considerations for limiting the time required for their involvement in my interviews and to respect the timing and the way they chose to participate (e.g. face-to-face or telephone interviews as convenient for them).

Specifically, I had different experiences and challenges in this regard during my fieldwork in China for Case 1 and Bangladesh for Case 2. In China, being a local citizen and speaking my native language – Mandarin – enabled me to communicate with my respondents easily and then save more of their time in taking my interviews. In addition, my familiarity with the local context and culture helped me better understand responses gathered and also reduce the risks of my interview being perceived as a disruption and intrusion into respondents’ work and life.

However, there were more challenges in Bangladesh as I was totally new to the local context and culture as an outsider, and my respondents and I communicated in a foreign language – English (most of my respondents speak English). I therefore took up more time of my interview participants than in China, as I needed to ask some more questions and clarifications in order to orient and familiarize myself with the context and background, and to make sure that I understood their responses properly. In this way, my presence and my research conduct in Case 2 in Bangladesh could be possibly more of an intrusion than in Case 1 in China.

Nonetheless, I observed low level of trust among some respondents when I was



conducting fieldwork in China. For example, one interviewed respondent from the supplier boundary partner group agreed to be interviewed but did not really reveal any particularly useful information due to concerns about confidentiality and information use, even though I repeatedly explained the research rationale and ethics, including assurance of anonymity and information use. In Case 2 in Bangladesh, respondents showed slightly higher level of motivation and trust in participating in my primary research as there was interest among project members to learn how the key capacity building element could be assessed through my research activities.

An additional ethical challenge during my fieldwork lay in the fact that many of my interview respondents themselves (e.g. those identified as boundary partner group members) in both case studies were the key project actors implementing the project and delivering services towards the achievement of desired project objectives/results. There might be potential issue of bias in that these respondents might have answered my evaluation-related interview questions (e.g. progress-marker-related questions) in ways that may make them or their performance 'look good' in attaining expected results. Also, there were 'gatekeeper' type of project actors (I define them as members in lead project team who allowed me access to the projects and to relevant project information) identified as my interview respondents. They also had an interest or stake in any forms of assessment of the achievement of project results and therefore may potentially risk framing answers to my project-effect-related questions in a biased way. To mitigate these above-mentioned potential problems, I needed to refer to and triangulate information and evidence from various different sources. For example, there were questions examining similar issues designed for different types of respondents for such triangulation purpose (see interview guides in appendices).

## **Chapter 5 Application of the Synthesis Model – Case 1**

With the guidance of the evaluation design and case study analysis protocol described in the previous chapter, in chapters 5 and 6, I present in more detail how I specifically applied the synthesis model in the two case study projects; unpack the evaluation function of the model through an analytical breakdown of its various key components; and conclude with an extended within-case analysis to further examine whether the use of the synthesis model adds distinctive value to each case study project, and exemplify the case-specific conditions under which I applied the model as reference for future usage.

Following the same structure, each chapter begins with a brief summary of the key features of the case study project. Second, it introduces the reconstructed project framework using the synthesis model and the process of defining some of its key elements. Third, it identifies and analyses various components of the project according to the synthesis model, following the sequence – actor analysis, results chain analysis, and evaluation mechanism analysis, as introduced in Section 4.2.5. Finally, it proceeds to an extended within-case analysis of the actor component of each project, a discussion linking and comparing the evaluation mechanism of the synthesis model with that of the LFA and the project's existing system, and an analysis of influencing factors in the case study project under the framework of the synthesis model.

### **5.1 Case 1**

#### **5.1.1 Key features of Case 1**

In the first case study, I applied the synthesis model to an EU-funded project on sustainable consumption with the ultimate goal of promoting resource efficiency and

environmentally friendly economic development in China. On the basis of the general project introduction in Chapter 4, Table 5.1 summarizes the key features of Case 1 in terms of a set of parameters relating to:

- Project type, duration, geographical location, funding source, and influencing/contextual factors
- Composition of key components
- LFA application (extent to which the project used the LFA)
- OM application (extent to which the project used OM)
- Overall existing evaluation system in the project
- Behavioural change and capacity building as key elements

**Table 5.1: Key features of Case 1**

<b>Key features</b>		<b>Case 1</b>
<b>Type of project</b>		Aid project on sustainable consumption
<b>Duration of project</b>		3 years
<b>Geographical location</b>		China (2 megacities)
<b>Funder</b>		EU
<b>Overall goal</b>		To promote resource efficiency and environmentally friendly economic development through mainstreaming individual sustainable consumption, and to improve the quality of life in the target area
<b>Actors</b>	<b>Lead implementation team</b>	BUCEA, NKU, IPPR, C2020, BTCA
	<b>Boundary partners</b>	4 groups: <ul style="list-style-type: none"> <li>• Local consumer associations</li> <li>• Representative retailers</li> <li>• Representative suppliers</li> <li>• Local authorities</li> </ul>
	<b>Beneficiaries</b>	Local consumers
<b>Outcome challenges</b>		4 sets of outcome challenges
<b>Progress markers</b>		4 sets of progress markers
<b>Behavioural change and capacity building as key elements</b>		Both elements can be seen in the project
<b>LFA application (extent to which the project used the LFA)</b>		Planning, monitoring and evaluation (partial usage)
<b>OM application (extent to which the project used OM)</b>		No OM elements adopted

<b>Evaluation system (existing)</b>	<ul style="list-style-type: none"> <li>• Internal evaluation through the LFA</li> <li>• Results-oriented monitoring</li> <li>• External evaluation using EU framework against criteria such as efficiency, effectiveness, impact and sustainability</li> </ul>
<b>Influencing/contextual factors</b>	<ul style="list-style-type: none"> <li>• Continuous government support of sustainable consumption policies</li> <li>• Local government support of sustainable consumption</li> <li>• Willingness of retailers and suppliers to jointly promote sustainable consumption</li> <li>• Motivation and willingness of local consumers to purchase green products</li> <li>• Green markets are created, small and medium-sized enterprises (SMEs) – suppliers – profit by supplying green products, and citizens increasingly buy green products</li> </ul>
<p><b>Note:</b> boundary partners, outcome challenges and progress markers are OM terms adopted and reformulated in the synthesis model, rather than from original project documents; NKU, IPPR, C2020, BTCA are used as proxy names of institutions involved in the lead implementation team.</p>	

(Source: BUCEA, 2012; 2013; 2015)

The following section unpacks the application process of the synthesis model in Case 1 in detail: the formulation process of a holistic project framework using the logic of the model and how each level of result, actors and indicators were defined in the project matrix.

### **5.1.2 Application of the synthesis model and formulation of a holistic project framework**

As Table 5.2 shows, with initially gathered secondary data, I formulated a holistic and detailed project framework using the logic of the synthesis model, including major project elements located both vertically and horizontally in the matrix: project results at various levels, corresponding indicators and groups of actors. In addition to following some of my general criteria for case study selection (as specified in the previous chapter), it can be seen from the project framework that Case 1 not only has a strong capacity building element (behavioural change expected as a key result of project objectives), but also has clearly and easily differentiated boundary partner groups. These project features and elements suggest that it is appropriate and

potentially valuable to apply the synthesis model rather than the LFA (or OM) alone to this case.

In this reconstructed project framework, there are seven rows and three columns accommodating the key elements of this case study project. This is mainly the version from my interpretation as a researcher and external independent evaluator with relatively limited time, resources and available data during fieldwork. In my application of the model, influencing and contextual factors were also noted (elaborated in Section 5.2.3 of this chapter) but are not directly included in the framework in the interests of presenting the key elements in a simple and clear manner.

**Table 5.2: Case 1 reformulation framework**

Reformulation framework with the synthesis model		
Objectives/results	Content/statement	
<b>Overall goal/impact</b> (beneficiary level)	Promote resource efficiency and environmentally friendly economic development in China through mainstreaming individual sustainable consumption, and improve quality of life in the target area	
<b>Project goal</b>	Foster and mainstream sustainable consumption patterns and behaviour through business networking of multi-stakeholders, and promotion of green product market by awareness raising, capacity building of consumer associations through intensive EU–China partnership and practices/experience exchange	
<b>Project outcome</b> (behavioural change in boundary partners)	<b>Boundary partner:</b> Local consumer associations	<b>Outcome Challenge:</b> Local consumer associations contribute to fostering and mainstreaming sustainable consumption patterns and behaviour through business networking with multi-stakeholders, and promoting green product market by awareness raising, capacity building, and effective information dissemination
		<b>Progress markers</b>
		<b>Expect to see:</b>
		1. Signing voluntary agreements (VAs) with larger retailers/supermarkets and SME suppliers
		2. Working with evidence-based communication campaigns often involving product tests, and putting pressure on suppliers and producers of harmful or non-sustainable products to change their supply or production approaches
		<b>Like to see</b>
		3. Disseminating information on product quality comparison tests to consumers
		4. Including sustainable consumption websites developed during this project into the websites of consumer associations of the two cities
		<b>Love to see</b>

			5. Raising awareness and strengthening capacity in helping mainstream citizens' sustainable consumption patterns and behaviour
			6. Transferring knowledge on best practices in sustainable consumption between the EU and China
	<b>Boundary partner:</b> local authorities	<b>Outcome Challenge:</b> In collaboration with other project partners, local authorities provide support to promote sustainable consumption through policy dialogue, setting up relevant requirements, and seeking networking and synergy with other ongoing programmes	<b>Progress markers</b>
			<b>Expect to see:</b>
			1. Keeping policy dialogue with project partners
			2. Seeking synergy between the project and related local programmes in the two cities
			<b>Like to see</b>
			3. Supporting awareness raising and capacity building in consumer associations etc.
	<b>Boundary partner:</b> representative retailers	<b>Outcome Challenge:</b> Retailers are encouraged to increase the percentage of green products on their shelves, transfer knowledge on sustainable product labels, and disseminate tools and skills to grow market demand for sustainable products	4. Raising citizen awareness of sustainable consumption through support to training schools and relevant courses
			<b>Love to see</b>
			5. Transferring knowledge on best practices in sustainable consumption between the EU and China
			<b>Progress markers</b>
			<b>Expect to see:</b>
			1. Signing voluntary agreements with local consumer associations
	<b>Boundary partner:</b> representative suppliers	<b>Outcome Challenge:</b> Through business networking with other project partners, SME suppliers are motivated to provide more green products and make their production chains more environmentally friendly from design to recycling through technical and managerial innovations	2. Signing green supply contracts with SME suppliers to help establish product sustainability criteria by introducing green labels and providing information on energy costs
			<b>Like to see</b>
			3. Giving priority to buying sustainable products from SME suppliers
			4. Increasing percentage of green products on shelves
			5. Making efforts to influence suppliers and promoting choice editing
			<b>Love to see</b>
	<b>Boundary partner:</b> representative suppliers	<b>Outcome Challenge:</b> Through business networking with other project partners, SME suppliers are motivated to provide more green products and make their production chains more environmentally friendly from design to recycling through technical and managerial innovations	6. Transferring knowledge on sustainable product labels and disseminating tools and skills to grow market demand for sustainable products
			<b>Progress markers</b>
			<b>Expect to see:</b>
			1. Signing voluntary agreements with local consumer associations
			2. Signing green supply contracts with target retailers
			<b>Like to see</b>
	<b>Boundary partner:</b> representative suppliers	<b>Outcome Challenge:</b> Through business networking with other project partners, SME suppliers are motivated to provide more green products and make their production chains more environmentally friendly from design to recycling through technical and managerial innovations	3. Promoting the development of green supply and procurement
			4. Taking innovative technical and managerial measures to produce more sustainable products
			5. Increasing the transparency of product information
			<b>Love to see</b>
			6. Taking efforts to make their production chains more environmentally friendly from design through to recycling
			<b>Output indicators:</b>
	<b>Output</b>	<ul style="list-style-type: none"> <li>• Supply of green products to target retailers increased by 10% in the first year and 15–20% in the third year, as compared to 2010 baseline</li> <li>• 20 VAs signed in the first year and 40-60 VAs signed in the third year in which green product standards should be met</li> <li>• 400 SME suppliers signed green product supply contracts in the first year and 1200–1500 SME suppliers signed green product supply contracts in the third year</li> <li>• Increase of Euro (EUR) 30 million for sustainable public procurement in participating public institutions and their network organizations at the end of this action, compared to 2010 baseline</li> <li>• By increasing green product procurement, saving 120 million tons of water, and 40</li> </ul>	<ul style="list-style-type: none"> <li>• Share of green products in total supplied products to target retailers in the baseline year of 2010 (%)</li> <li>• Number of retailers signing VAs with local consumer associations in which green product standards should be met</li> <li>• Number of SMEs signing green product supply contracts with the target retailers</li> <li>• Expenditure on sustainable public procurement of participating public institutes and their network organizations</li> <li>• Reduced energy consumption, reduced CO2 emission, reduced water consumption, and reduced waste generated, as compared to 2010 baseline</li> </ul>

	thousand tons of waste reduced, as compared to 2010 baseline	
<b>Project activities</b> (project team)	<ul style="list-style-type: none"> <li>• Inception and project management</li> <li>• Awareness raising and capacity building in target region</li> <li>• Implementation and support</li> <li>• Monitoring and enforcement</li> <li>• Up-scaling and policy dialogue</li> </ul>	Strategy map/journal (concept in OM for monitoring, an optional choice)

As the above table shows, at a vertical level, and as shown by the first and second columns, there are two broad types of element: 1) distinctive and explicit groups/categories of actor such as the project team, various groups of boundary partners with whom the project team interacts directly, and beneficiary level actors. As indicated by the authors of OM (Earl et al., 2001), boundary partners can be nested and may have their own boundary partners. The notion of ‘boundary partner’ is therefore relative and depends on the reference point. This means the list of actors can be longer in a more complex project involving many more different institutions and more complicated cooperation mechanisms (more specific analysis of the actor component of this case study project is presented in the next section); and 2) a hierarchy of levels of results following the general structure of a conventional approach such as the LFA and in a sequence of activities, output, outcome, goal and impact (although the connotation and definition of some levels of results are different).

Output-level results basically refer to a similar concept as that with the LFA. On the other hand, outcome-level results are defined in OM terms as behavioural changes in project boundary partners in particular, and goal- and impact-level results are related to beneficiary-level changes. As the third column of the table shows, this framework includes performance indicators corresponding to each level of result which combine both quantitative and qualitative parameters.

When formulating the above project framework, I referred mainly to secondary sources of data from the project team and the donor side. I relied heavily on the original project logframe and introductory project information in formulating output- and impact-level results as they could easily be transferred directly from the original logframe to the matrix of the synthesis model due to similar concepts and related groups of actors. However, it took me considerably more time to reconstruct middle-level results – outcomes and goals – using the synthesis model, due to the difference between my definition and that of the original project framework in which actors were relatively more abstract and there was only one type of middle-level result named as ‘specific objective’ (BUCEA, 2012).

In contrast, in applying the synthesis model, I needed first to make the actor component explicit by identifying and distinguishing relevant actors linked to different levels of results – boundary partners in relation to outcome and specific beneficiaries in relation to goal. I then needed to extract relevant data from multiple sources of available project documents. This information extraction and reinterpretation process was time consuming as the boundary between actors and these two levels of results was blurred in the original project framework. For example, in the original project information sheet, the following five groups of actors were all listed under the category of target groups (BUCEA, 2012; SWITCH-Asia, 2015a):

- 1) Local consumer associations in Beijing and Tianjin
- 2) Retailers
- 3) SME suppliers
- 4) Local authorities
- 5) Consumers in the megacities of Beijing and Tianjin



On the basis of a detailed review of all the available secondary data – those in the original project logframe in particular – I separated this list of actors into two groups: boundary partners (the first four groups) and beneficiaries (the fifth group). I then differentiated their respective project activities, roles and responsibilities, and set targets with indicators according to gathered project data.

During the process of reformulating the project framework using the synthesis model, other than the effort taken in clarifying the actor component, defining and interpreting another two elements – outcome challenge and progress markers – were also time consuming. This was mainly due to the fact that the synthesis model adopts the OM concept of ‘outcome’ together with two related elements: ‘outcome challenge’ and ‘progress markers’. This was a new element in Case 1 and there was no corresponding data directly transferrable from the original project framework to the synthesis model, so I needed to construct the ‘outcomes’ from scratch.

I did this by first identifying the relevant actors – boundary partners linked to outcome challenges and progress markers. I revisited all the available project information (e.g. logframe, narrative reports) to roughly categorize and code each boundary partner group against outcome challenges and progress markers. I then summarized each outcome challenge with one sentence to fit in the matrix (as shown in the second column of Table 5.2). The determination of progress makers and three related progressive ladders ‘expect to see’, ‘like to see’ and ‘love to see’ involved additional subjective construction according to my knowledge and experience of OM concepts. I further coded the information gathered on progress makers (mainly related to the activities and behaviours of boundary partners), and assigned differing degrees

of progression from minor behavioural/activity changes at an early stage to more profound influence and transformation at later stages.

As discussed in Chapter 4, I used outcome challenges and progress markers that I constructed in the project reformulation framework at this stage as a guideline during my primary data collection. I determined these two elements based on limited but key project documents. Most of them proved relevant and were validated during my primary research through interviews with various project actors.

For example, the following quotations represent source data that I referred to when constructing some of the progress markers listed in Table 5.2 (BUCEA, 2012):

The project aims at improving the supply chain by facilitating voluntary agreements between larger retailers (supermarkets) and consumer associations. The signed agreements encourage retailers to increase the percentage of green products on their shelves. The project engages retailers, encouraging them to transfer knowledge on sustainable product labels and to disseminate tools and skills to grow market demand for sustainable products. The project uses both ‘pull’ and ‘push’ strategies and addresses producers, suppliers and consumers.

Green supply contracts are signed between retailers and SME suppliers with support from the project. The contracts support retailers in establishing the sustainability criteria for products by introducing green labels and providing information on energy costs throughout the life cycle. Local consumer associations disseminate information including that of product quality comparison tests to consumers. By doing so, the demand for green products increases and SME suppliers are motivated to sign more green supply contracts.

With reference to these data, I inferred some progress makers related to retailers at two levels: signing voluntary agreements with local consumer associations and green supply contracts with SME suppliers at the ‘expect to see’ level; and transferring knowledge on sustainable product labels and disseminating tools and skills to grow market demand for sustainable products at the ‘love to see’ level.

I began this first stage of secondary data analysis before primary data collection so

that I could formulate the project framework using the synthesis model to provide an evaluation guide with targeted key project partners whose behavioural changes were to be traced and corresponding primary data collected during empirical research. In the following sections, I focus on more in-depth, second-level analysis of three aspects with the purpose of unpacking various components and evaluation instruments of the synthesis model: actor analysis, results chain analysis, and evaluation mechanism analysis. As specified in Chapter 4 on the overall data analysis strategy, these three aspects aim to answer key evaluation questions in terms of **who** (are the major actors and focus of evaluation design from an actor-centred perspective), **what** (specific results/changes are being examined) and **how** (to evaluate these changes, i.e. the process and means deployed for guiding and operationalizing evaluation data collection and interpretation).

In addition, by taking an actor-oriented approach to case study analysis, I seek to 1) explore the change process from micro to macro level with a focus on major actors involved at different stages of change; 2) uncover the evaluation function of the synthesis model; and 3) examine the added value of the model by explicitly bringing in the actor component.

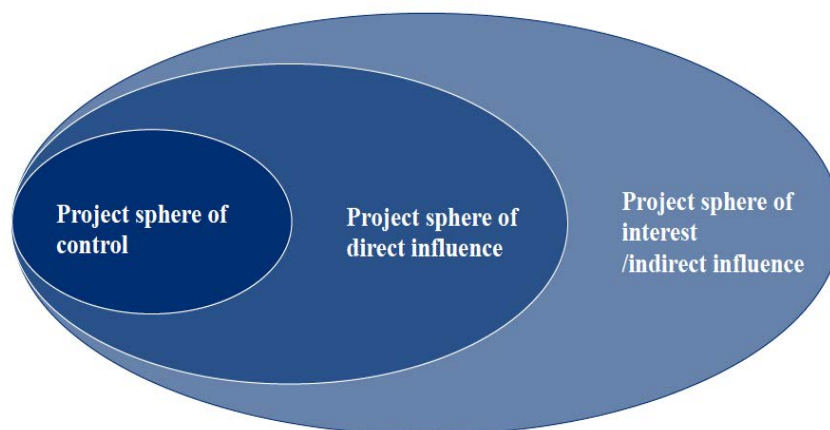
### **5.1.3 Actor analysis**

As noted previously, ‘boundary partner’ is one of the key innovative concepts imported from OM to the synthesis model. This not only indicates a clear actor-oriented perspective, but also implies that there are different categories of actors located in different domains with their respective roles and responsibilities in a particular project. As the Case 1 project reformulation matrix in Table 5.2 shows, I identified three groups of actors involved in this project as follows:

- Lead project implementation team
- Boundary partners – local consumer associations, representative retailers, representative suppliers, and local authorities
- Beneficiaries (local consumers in the two targeted cities)

As argued in Chapter 3, such a clear division of actors seems to be implicit in a conventional approach such as the LFA. However, in clearly identifying and defining boundary partners, the limit of the project's sphere of influence is acknowledged (as noted in Section 3.2.2 of Chapter 3). Figure 5.1, shows three domains of project influence represented by three ellipses in various degree of blue shading which gradually fades from the first to the third ellipse, illustrating correspondingly decreasing project influence in each domain (Earl et al., 2001). I adopted this sphere of influence representation to facilitate both actor analysis and the subsequent results chain analysis in each case study.

**Figure 5.1: Project sphere of influence**

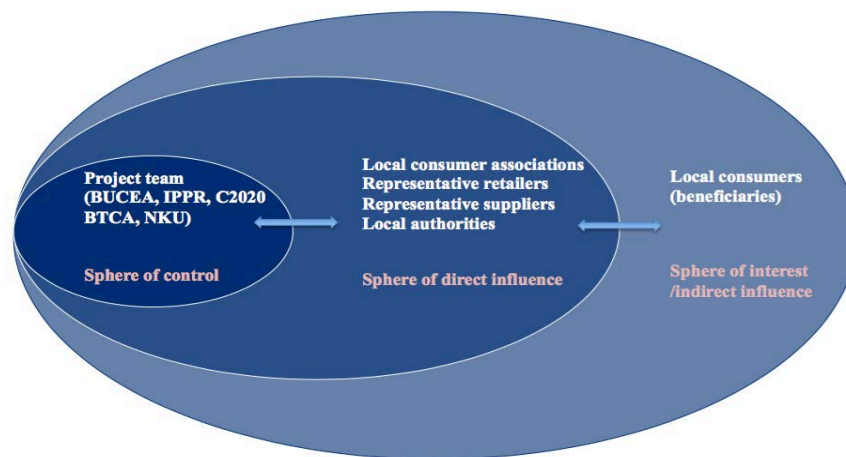


(Adapted from Earl et al., 2001; Hearn and Ambrose, 2013)

I used this 'sphere of influence' notion to guide my identification of boundaries and boundary partners (many of whom were my primary data sources), and to locate them

in different project domains. This helps illustrate the utility of the synthesis model in terms of its actor component – adding the additional actor strand of the theory of change in contrast to the single results chain related theory of change in the LFA alone. Using the sphere of influence concept as an actor analysis tool, I constructed Figure 5.2 below to indicate the chain of actors, their respective project domains, and their relationships.

**Figure 5.2: Actor strand of the theory of change in Case 1**



As noted in the previous section, boundary partners can be nested and the reference point – whose boundary partners are to be targeted – is vital in drawing boundaries among various actors along the actor chain. In both case study projects, the lead project team (who was under direct supervision of the donor agency and directly implemented the project in the field) was my reference point and their boundary partners were identified and focused on during the empirical research process.

In Case 1, a lead team comprising five members implemented the project through mobilization and operationalization of resources, and delivery of corresponding services and activities. I therefore located them in the sphere of control, as shown in Figure 5.2. I based this decision on both secondary project information and some

primary data gathered from interviews with relevant project team members. With this lead team as a reference point (meaning that their boundary partners were to be targeted), I identified four groups of boundary partners – local consumer associations, representative retailers, representative suppliers, and local authorities – and situated them in the sphere of direct influence. The lead project team had to work with these boundary partners directly in order to reach local consumers (beneficiaries) who remained beyond the sphere of direct project influence. For example, the project team worked directly on raising awareness and strengthening the capacity of local consumer associations so that they could then communicate with local citizens directly and help foster mainstream sustainable consumption patterns and behaviours among them.

Distinguishing the layers of actors involved in the project was the first step in my application of the synthesis model. These actors were the driving force of change and their roles and responsibilities needed to be clarified if changes were to be achieved through project endeavour. Moreover, as the bidirectional arrows in Figure 5.2 indicate, according to the synthesis model, relationships between actors were interactional. Thus, actors in each domain exerted influence on and effected changes in other group of actors, but were also affected by them. This accords with Davies (2004, p.105) in the suggestion that various types of actors at different stages of the project should be clearly defined, particularly in the sequence that ‘they were [affected] by and affect[ed] others’. Through their implementation of project activities and service delivery, the project team produced certain outputs so that the capacity of these groups of boundary partners could be built and changes in their behaviours effected. They then continued to influence or contribute to further changes at beneficiary level.

Such actor identification and analysis helped in clarifying the division of roles and responsibilities of each group of actors, and the relationships among them. It also helped in the formulation of an actor strand of the project theory of change, providing clear guidance for identifying my primary data sources – those who had changed and from whom evidence or stories of change could be collected. For example, by constructing the actor chain shown in Figure 5.2, I perceived that the project team would be my best primary source of output-level results, and the four boundary partners best placed for understanding changes related to behaviours and capacity at outcome level.

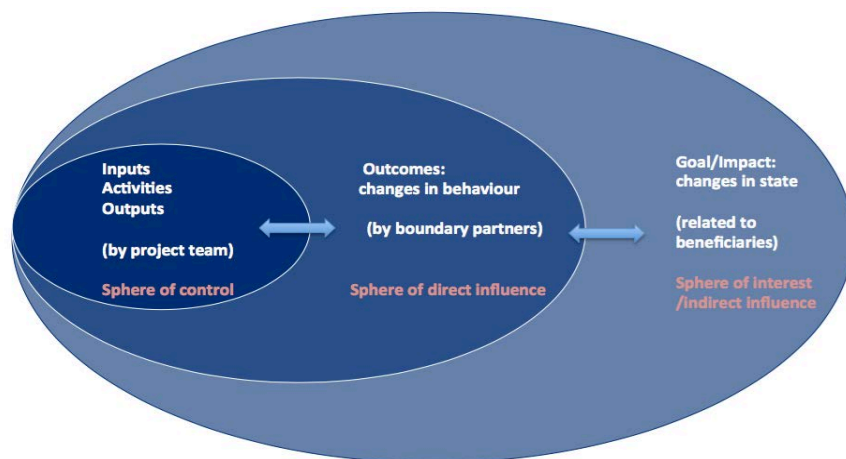
Furthermore, making explicit respective roles and interconnected relationships among actors could enable a more nuanced and better-focused process of review, reflection and improvement of such relationships (although this was not fully explored in my empirical research as the ‘improvement’ aspect requires continuous involvement with project partners over a longer period). For example, one group of actors in the lead project team in the sphere of control, namely, local consumer associations, also belonged to the group of boundary partners in the next domain – the sphere of direct influence – as shown in Figure 5.2. They were not only key players in the lead project team implementing major project activities, but also intermediary actors whose capacity was strengthened and behaviours changed in order to deliver better services to local consumers at beneficiary level. I needed to take into account their dual role and relationships with other actors in my evaluation design and primary data collection. Therefore, I designed two sets of interview questions (see appendices 5 and 9) for this particular group of actors which respectively focused on their role as a member of the lead project team and as a group of boundary partners.

#### 5.1.4 Results chain analysis

Having identified and defined the various groups of actors as required under the synthesis model, and having explored their positioning using ‘sphere of influence analysis’, I could then refine the focus of the evaluation I was going to carry out and identify whose perspectives on change I should gather and analyse. These actors included the lead project team, their boundary partners, and local consumers in the two targeted cities as the ultimate beneficiaries of the project. Other than the latter, I collected primary data from all other actor groups on their perspectives on actual change.

I then developed a results chain analysis to identify the types of change or results (linked to these actors already identified) I needed to look for and explore. As Figure 5.3 below indicates, I adopted the ‘sphere of influence’ tool and fleshed out the results aspects of the theory of change (as distinct from the actor aspects I outlined in the previous section) to help explicitly illustrate the distinctive project domain of each level of result and its corresponding actors according to the logic of the synthesis model.

**Figure 5.3: Results strand of the theory of change in Case 1**





As Figure 5.3 shows, the synthesis model results chain runs from project outputs to outcomes and goal and impact-level results. Thus, utilizing the sphere of influence concept, the theory of change I formulated on the basis of this results chain reveals not only the domain of each level of result with its corresponding groups of actors at each stage of the project, but also their nonlinear and interactional relationships. Along this results chain, outputs, goal and impact-level results have a largely similar rationale to those of traditional evaluation approaches such as the LFA or RBM. However, the synthesis model brings in the OM concept of outcome as behavioural change in boundary partners. The outcomes in the reconstructed results chain for Case 1 therefore differed from those in the original project documentation.

According to the theory of change of the synthesis model, the inclusion of an additional level of results related to behavioural change in intermediate actors makes the results chain multi-dimensional and more complete than that of the LFA or OM alone. It bridges the gap between outputs and their effects along the results chain with distinctive groups of actors, and displays the change process and pathway more explicitly. It also helps identify the area in which project success or failure occurs and the actors to whom these results relate, so that learning can be achieved and improvements possibly made.

In Case 1, the project team generally had direct control over output-level results. They then extended their influence further along the results chain to beneficiaries through working with the four groups of boundary partners. Changes arising among the latter were crucial in effecting and sustaining changes or results at goal and impact level, and in helping to explain how and why some higher-level results occurred while others did not through focusing on micro-level cumulative changes. Additionally,

embedded in this results chain is the ‘relative influence’ notion (Smutylo, 2001, p.5), which states that the influence of the project decreases along the results chain and the owners of change move from project team to four groups of boundary partners to local consumers in this case.

### **5.1.5 Evaluation mechanism analysis**

I conducted an actor analysis to identify the focus of my evaluation in terms of whom to gather change stories from, and a results chain analysis to define the types of change and results to focus on when building the evidence chain and storylines. In this section, I further unpack the application of the synthesis model and answer the question of how to measure and make sense of these changes through an evaluation mechanism analysis. As briefly mentioned in Chapter 4, I use the term ‘evaluation mechanism’ to refer to the process and means involved and deployed in guiding and operationalizing evaluation data collection and interpretation. Specifically, I look into 1) the evaluation instruments accommodated in the synthesis model, and 2) the options and possibilities for analysing and making sense of the evaluation data acquired through these instruments.

First, at an operational level, I relied heavily on two instruments during my evaluation: various levels of results linked to their respective groups of actors, and the corresponding indicators. This procedure is similar to evaluation using the LFA and OM as both approaches utilize targets and results and related indicators in operationalizing their evaluations. For example, as seen from the reformulated project framework for Case 1 in Table 5.2, there were output indicators, outcome challenge/progress markers, goal indicators, and impact indicators linked respectively to output-, outcome-, goal- and impact-level results. I then used these indicators as

assessment parameters and guidelines for collecting evaluation data from both primary sources on progress-marker-related behavioural change in boundary partners, and secondary sources on results data in relation to output (project team), goal and impact (beneficiaries) indicators.

Second, another component in evaluation mechanism analysis is to determine the possible means of analysing and interpreting gathered data to tell a performance story, form an evidence chain, and then apply it to serve a specific evaluation purpose. In this way, I also sought to reveal whether the evaluation data collected from applying the synthesis model could be easily analysed and aggregated for further utilization. This was crucial as a sound evaluation tool should enable users to collect adequate and quality data which should be aggregatable and usable.

There are various options for analysing evaluation data: one can analyse the final results with reference to a baseline parameter or situation if particularly interested in before/after-type impact assessment; or one can apply a grounded theory perspective if particularly interested in capturing unintended or unforeseen dimensions of the process (Thomas, 2006; Van Ongevalle, et al., 2009). These principles are applicable to the evaluation data obtained with the synthesis model and, as mentioned in Chapter 4, I conducted analysis of my collected primary data by applying elements of both options.

For example, I analysed primary data that I had gathered specifically on behavioural change in identified boundary partners in Case 1 through identifying, clustering and categorizing similar behaviour changes, and then displayed the end product of the analysis in a matrix with a narrative description (see example data analysis matrixes

in tables 5.3 and 5.4). There are other ways of analysing data collected through progress markers, such as quantification of qualitative data by percentage, ratio, scoring, numerical value, etc. (Armstrong et al., 2000). I did not explore all these means of data analysis in this study as the ultimate research objective was not to compile a comprehensive evaluation report on project performance, but to exemplify how the synthesis model was adopted in my case studies and draw implications on its utility. As my evaluation of Case 1 utilized both quantitative and qualitative measures and indicators, I applied some of the multi-level data analysis strategies applicable for mixed methods evaluation including parallel mixed methods data analysis and multi-level mixed methods analysis<sup>35</sup>. Some illustrative examples are given in the following sections.

Means of data analysis and the style of presentation of its results largely depend on the objectives of the evaluation and the design of the data collection process. For example, in my interview questions during primary data collection, numerical values were not included. I then mainly used narrative descriptions and illustration matrixes as feasible options to present areas of achievement of progress markers among selected boundary partners.



















Tables 5.3 and 5.4 show results from the analysis of some primary data I collected during fieldwork. I constructed both tables from responses by representatives of two different groups of boundary partners to my progress-marker-related interview questions. The percentage rate of progress marker achievement of representative

---


<sup>35</sup> Parallel mixed methods data analysis involves two separate analysis processes: QUANT data are analysed using conventional QUANT methods (such as frequency tables, cross-tables, regression analysis, etc.) while a separate analysis of QUAL data is conducted using QUAL methods such as content analysis. The findings of the two sets of analysis are then compared. Multi-level mixed methods analysis adopts QUANT and QUAL analysis techniques at different levels of a multi-level evaluation design (Teddlie and Tashakkori, 2009).

retailers can be seen in Table 5.3: for example, there was a 50% achievement rate for ‘expect to see’ progress marker 1 across selected groups of representative retailers, while it was 100% for progress marker 2. This matrix also helps tell the performance story of individual members of a boundary partner group and shows how their performance in achieving the same progress marker compares with that of other actors. On the one hand, this addresses the aim of reporting for upward accountability and, on the other, provides an opportunity for tracing the reason for varied performance in order to realize the learning and improvement objective of evaluation.

**Table 5.3: Boundary partners – representative retailers’ achievement of progress markers<sup>36</sup>**


















Achievement of progress markers by boundary partners (representative retailers)				
Progress markers	Retailer 1	Retailer 2	Retailer 3	Retailer 4
Expect to see				
1				
2				
Like to see				
3				
4				
5				
Love to see				
6				

Other than interpreting evaluation data to tell the performance story of boundary partners, other dimensions can also be explored such as examining the design of progress markers. For example, Table 5.4 shows achievement for another group of boundary partners, local consumer associations, four out of five of whom did not


<sup>36</sup> The progress markers in Table 5.3 correspond to the row in Table 5.2 which contains the boundary partner group – representative retailers. The symbol  denotes completion of the relevant progress marker as reported by the boundary partner interviewed.

complete ‘expect to see’ progress marker 1. This could mean either a flaw in progress marker design or underperformance of particular boundary partners, or, in other words, design failure or implementation failure. Which of these explanations applies can then be investigated by gathering further information from relevant boundary partners, and by reviewing the intervention logic and links between outcome-level results and other levels of results along the chain. Such further probing can also help identify internal and external hindering factors.

**Table 5.4: Boundary partners – local consumer associations’ achievement of progress markers<sup>37</sup>**

Achievement of progress markers by boundary partners (Local consumer associations)					
Progress markers	LCA 1	LCA 2	LCA 3	LCA 4	LCA 5
Expect to see					
1					
2					
Like to see					
3					
4					
Love to see					
5					
6					

The achievement of progress marker 6 (a higher-level ‘love to see’ indicator on transferring knowledge on best practices in sustainable consumption between the EU and China, as detailed in Table 5.2) was also low, only two out of five selected groups completing it. The reason for low achievement of this progress marker was not likely to have been due to flawed design as it was one of the key outcome-level indicators

<sup>37</sup> The progress markers in Table 5.4 correspond to the row in Table 5.2 that contains the boundary partner group – local consumer associations. LCA here refers to individual local consumer association involved in my interview. The symbol  denotes the completion of relevant progress markers as reported by the boundary partner interviewed.

and strongly linked to the project goal. Interviews with respondents from these local consumer associations revealed implementation challenges in achieving this progress marker:

Due to considerable differences existing in specific conditions and the status quo of consumption patterns between China and EU countries, some particular China-specific consumption problems don't apply to EU countries. Under these circumstances, learning from the achievements and successful experiences from EU counterparts on sustainable consumption proved to be limited. The cooperation between China and EU partners should not be just confined to limited opportunities for joint seminars and discussions; the two parties should have had more in-depth exchanges and carried out joint field visits and research to discover real-time challenges and difficulties, and to explore tailored and effective solutions. (Boundary partner interview 6, November 2014, Beijing, China)

In general, results gathered through the combination of evaluation instruments and mechanisms of data analysis and interpretation incorporated in the synthesis model can help build an actual results chain supported by evidence and storylines from minor behavioural change to macro-level transformation in corresponding groups of actors.

## **5.2 Uncovering the distinctive value in applying the synthesis model**

The above section presented a detailed analysis of a series of components of Case 1 framed under the synthesis model to systematically unpack the application process and evaluation function of the model with empirical data and examples of usage. In this section, I extend the discussion to further examine whether the synthesis model adds value beyond that which the LFA or other approaches originally employed by this project were able to deliver on their own. I also seek to generate some implications for future use of the model in contexts similar to that of Case 1.

The following subsections include: 1) further discussion on the actor component of

the synthesis model; 2) linking and comparing the evaluation mechanism of the synthesis model with that of the LFA and the existing evaluation system in the project; and 3) exploration of the influencing factors at play to further contextualize the usage of the model and some case-specific conditions under which it can be applied.

### **5.2.1 An extended discussion of the actor component in Case 1**

As noted earlier in this chapter, the synthesis model explicitly incorporates an actor strand of the theory of change which indicates the actor structure, project domains of various groups of actors, and their non-linear and interactional relationships. In addition to determining who these major actors are and their positioning in the project cycle, clarity on their diverse characteristics is also important. ‘Messy partnerships’ (Guijt, 2008), as mentioned in Chapter 2, might be formed due to differences among actors in terms of their governance and cultural features, mandate, capacities, and commitment to collective objectives, etc.

Such messiness may exist not only across different categories of actors but also within the same group of actors. For example, along the actor chain of Case 1, the first actor group – the lead project team – comprised five member institutions ranging from universities to government and civil society organizations and international research institutions. The second group – boundary partners – included four member groups from a variety of sectors with diverse roles – local authorities, local businesses, and civil society organizations. Their different governance structures, capacities, and commitments to the project vision all tended to add to the complexity and messiness of partnerships and relationships.



As the actor component of the synthesis model unfolds, it affords greater clarity and the opportunity to understand individual characteristics, needs and interconnections. This in turn means that project actors can allow for and manage such complexity in their approach to implementation. This is a distinctive advantage compared with the LFA or OM. As reviewed in Chapter 3, the actor component of the LFA is comparatively abstract and the project theory of change is mainly articulated through a single results chain. Therefore, some characteristics of individual actor groups tend to be neglected and there is a potentially higher probability of forming messier partnerships. In OM, although there is a clear actor thread in building the theory of change, the chains of actors and results are relatively less complete compared with the synthesis model as they focus more on boundary partners and the contribution of their behavioural changes.

### **5.2.2 Linking and comparing the synthesis model with the LFA and the existing evaluation framework**

As Table 5.1 shows, in Case 1, the LFA played a key role in various project stages including internal evaluation. The project originally adopted the LFA from the planning stage and formulated a logframe matrix against which some results were measured and assessed by the project side. For example, in nearly all collected annual review reports on this project, there were clear indications that the logframe had been consulted in results assessment, as exemplified by the following quotation from the review questions to the lead project team in the donor-required standard report template:

What is your assessment of the results of the Action so far? Include observations on the performance and the achievement of outputs, outcomes and impact in relation to specific and overall objectives, and whether the Action has had any unforeseen positive or negative results (please quantify where possible; refer to Logframe Indicators). (BUCEA, 2012; 2013; 2015)

The project team reported back with achieved results against relevant indicators in the logframe as follows:

**Activity** A3.1: Implementation of VAs and MoUs [Memorandum of Understanding] (SC implementation) and post-ROM action plan task T2: Developing Green Supply Best Practices

The activity A3.1 is implemented in combination with the post-ROM action plan task T2. During the implementation of MoUs signed in this action, the action partners have visited the pilot retailers and have had intensive discussions. During the process, BJCA [Beijing consumer association] and TJCA [Tianjin consumer association], with the support of other partners, have selected 4 best practices in Beijing and Tianjin (it was planned in the post-ROM action plan task T2 to select 2 best practices). The 4 best practices created by this action are being disseminated to other retailers.

**Results** achieved:

- 1) Four best practices reports
- 2) Disseminated and presented best practices (BUCEA, 2013)

There were also other evaluation mechanisms in place such as ROM and donor-required external evaluation largely using OECD evaluation criteria. Relevant results are not available but an interview with a respondent from the donor side revealed some of the reality in evaluating this and similar projects funded by the EU:

Two lines of reporting: one is on budget reports on how the money has been spent. For such a project, there is no mechanism in place to ensure that the impacts are measured and monitored as strictly as financial aspects. However, this is also changing: the EU is now implementing a new results-based monitoring framework, not for Switch Asia projects yet, but for the future. This relates to the SDGs: there has been a call for a data revolution and sustainable development goals. There will be a lot more work and a lot of indicators on sustainable development. To report on this you need a lot of data, especially quantitative data. There will be a big push in the international development world to deliver this kind of data, so the EU is then of course in this overall framework. (Donor interview 2, January 2015, Beijing, China)

This quotation reflects the need for an evaluation mechanism to measure project results in addition to reporting for financial accountability. For example, the use of the LFA in internal evaluation followed a linear activity–result reporting model which did not sufficiently reflect the key change process from outputs to goal-level results and

thus had weak explanatory power. The synthesis model brought in an additional dimension in measuring project results with a holistic evaluation mechanism which functioned through a series of performance indicators related to each level of results and the corresponding actors. This is structurally similar to the evaluation system of the LFA but has two major differences: 1) an explicit actor division and actor strand of the theory of change; and 2) an additional level of results related to behavioural change in intermediate actors, which makes the results chain more complete and systematic than that of the LFA.

### **5.2.3 Analysis of influencing contextual factors in Case 1**

As discussed in Chapter 2, development interventions represent the process of external factors entering the existing context and becoming internalized in the local system. It is critical that evaluation reflects this process and considers how various contextual factors influence the changes it assesses. In addition, further study of context helps generate ‘under-similar-conditions’ generalizations. The synthesis model accommodates the consideration of influencing and contextual factors, as discussed in Chapter 3 (although this could potentially be presented in the form of assumptions, it is not reflected directly in the reformulated project framework in Table 5.2 to simplify the matrix). This section elaborates on these factors as they were manifested in Case 1 to illustrate their influence upon and relationship with different domains of results; and to show the case-specific conditions under which I applied the synthesis model to gather implications for further ‘under-similar-conditions’ projects.

With reference to the logic of the synthesis model and information from the original project logframe (e.g. the column of assumptions) and other relevant documentation, I constructed Figure 5.4 below to illustrate these major influencing factors in

connection with various levels of results in Case 1.

**Figure 5.4: Analysis of influencing factors in Case 1**

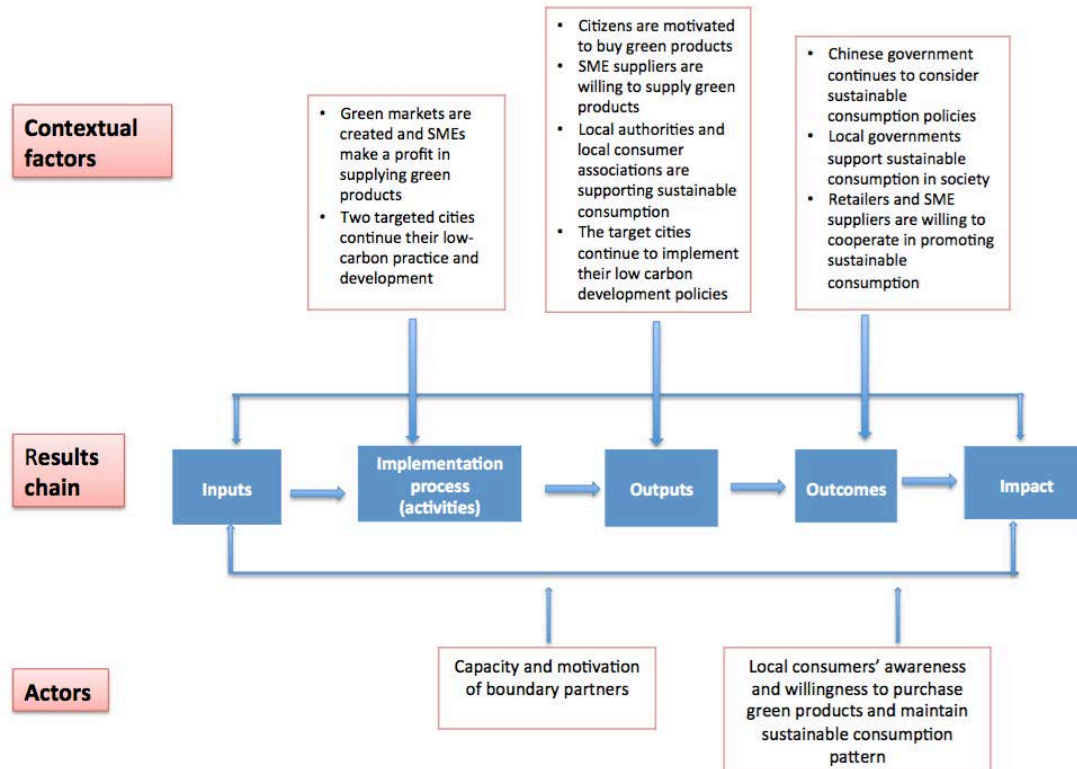


Figure 5.4 seeks to represent in a more explicit way influencing (including contextual) factors related to different levels of results along the results chain, as each stage of the project might require different conditions and encounter different circumstances, or some factors might affect certain results more than others. For example, as seen from the contextual factors in Figure 5.4, the availability of a market for green products with profit-making potential would be necessary if suppliers were to engage in providing such products – this is critical with regard to implementing project strategies and achieving expected outputs. However, in terms of outcome-level results, the policy environment would be a key influencing factor.

Under the framework of the synthesis model, the characteristics and capacity of

relevant actors are also regarded as key influencing factors. For example, without certain awareness, and the willingness and capacity to support and purchase green products on the part of consumers, a sustainable consumption pattern cannot be sustained. In the synthesis model, these contextual factors interwove directly or indirectly with other key components to effect and contribute to change at all levels.

## Chapter 6 Application of the Synthesis Model – Case 2

This chapter elaborates in detail the application of the synthesis model in Case 2 and follows the same case study analysis protocol, strategy and chapter structure as that for Case 1 in Chapter 5. This chapter comprises a comprehensive project reformulation framework using the synthesis model, and a series of analyses which unpack the evaluation function of the model through the breakdown of its key components: actors, results chain, and evaluation mechanism. The chapter then concludes with an extended within-case analysis to further explore the value added by applying the synthesis model in comparison with the LFA and other approaches originally adopted in the project, and to highlight some case-specific conditions under which the model was applied.

### 6.1 Case 2

#### 6.1.1 Key features of Case 2

In Case 2, I applied the synthesis model to a DFID-led programme in Bangladesh with the overall goal of achieving pro-poor economic growth to increase income and employment. The key features of this project are summarized in Table 6.1. Although it differs from Case 1 in respect of funding source, project sector and scale, and targeted geographical location and population, they share some common features such as a strong capacity building element and similar usage of the LFA as the major planning, monitoring and evaluation tool.

**Table 6.1: Key features of Case 2**

Key features	Case 2
Type of project	Aid project on poverty reduction through micro finance
Duration of project	7 years
Geographical location	Bangladesh (across northern, south-western and north-eastern regions)
Funder	DFID
Overall goal	Pro-poor economic growth to increase income and employment,

		especially for women, and to promote a sustainable micro-finance sector that offers greater access to and usage of diversified financial services for the poor.
<b>Actors</b>	<b>Lead implementation team</b>	PKSF
	<b>Boundary partners</b>	1 group: 24 partner organizations
	<b>Beneficiaries</b>	Ultra-poor households in targeted areas
<b>Outcome challenges</b>		1 outcome challenge for the whole boundary partner group
<b>Progress markers</b>		1 set of progress markers
<b>Capacity building as key element</b>		A central element in the project
<b>LFA application (extent to which the project used the LFA)</b>		Planning, monitoring and evaluation (partial usage)
<b>OM application (extent to which the project used OM)</b>		No OM elements adopted
<b>Evaluation system (existing)</b>		<ul style="list-style-type: none"> <li>• Internal evaluation through results-based monitoring with the LFA as the major tool</li> <li>• Longitudinal impact studies conducted by external institutions (I got access to two annual impact study reports)</li> </ul>
<b>Influencing/contextual factors</b>		<ul style="list-style-type: none"> <li>• Non-financial services (literacy, agricultural extension, business development, etc.) to complement financial services are available</li> <li>• No major external shocks to the economy of Bangladesh (macro-economic stability maintained)</li> <li>• No major natural disasters</li> <li>• Policy environment for the NGO sector remains favourable</li> </ul>
<b>Note:</b> boundary partners, outcome challenges and progress markers are OM terms adopted and reformulated in the synthesis model, rather than from original project documents.		

(Sources: Khalily et al., 2013; PKSF, 2013)

The following section unpacks the application process of the synthesis model in Case 2 with the focus on the formulation and reconstruction of the holistic project framework according to the logic of the model, and the breakdown and analysis of various components of the framework.

### 6.1.2 Application of the synthesis model and formulation of a holistic project framework

As with Case 1, before going to the field for primary data collection, I conducted

initial study and analysis of some available secondary data to reconstruct the project framework using the logic of the synthesis model. These major elements of the reformulated matrix are presented in Table 6.2, including project results at various levels, and corresponding indicators and actors.

**Table 6.2: Case 2 reformulation framework**

Reformulation framework with the synthesis model		
Objectives/results	Content/statement	
<b>Overall goal/impact</b> (beneficiary level)	Pro-poor economic growth to increase income and employment, especially for women	
<b>Project goal</b>	Sustainable micro-finance sector in Bangladesh that offers greater access to and usage of diversified financial services for the poor	
<b>Project outcome</b> (behavioural change in boundary partners)	<b>Boundary partners</b> (partner organizations)	<b>Indicators</b>
		<b>Impact indicators:</b>
		<ul style="list-style-type: none"> <li>Seasonal Vulnerability Reduction Index improves by %</li> <li>Access scale score for food security during period of seasonal vulnerability improves by % by 2014</li> </ul>
		<b>Goal indicators:</b>
		<ul style="list-style-type: none"> <li>85% of ultra-poor households have access to formal financial services</li> <li>90% of households increase monthly income by 50% throughout the year; equity in IGAs (income generating activities) increases by 50% among 75% of borrower households</li> <li>No. of working days increases by %</li> <li>85% of members make repayments from income; IGA scale increases and seasonal effects overcome</li> <li>Progression to higher loans in 70% of cases to expand IGAs</li> </ul>
		<b>Progress markers</b>
		<b>Expect to see partner organizations</b>
		1. Selecting appropriate beneficiaries, developing profiles of them, forming groups, introducing a flexible saving system, and holding issues-based discussions in group meetings
		2. Determining IGA-based credit needs, providing flexible credit facilities, and following up on the use of loans and debt situations of members
		3. Assessing loan needs of IGA loanees during the season of vulnerability and providing emergency loans
		<b>Like to see partner organizations</b>
		4. Identifying training needs for viable IGAs of selected members, providing practice-oriented IGA training (skills, market linkage), and establishing linkage between members and markets
		5. Assessing technical service needs of IGA practitioners and providing/arranging necessary vaccinations
		6. Preselecting suitable trainees for selected informal vocational training and following up on provision of training
		7. Undertaking door-to-door primary health care services, delivering selected medicines, providing referral to specialized and high-level government organization, NGO and private facilities
		8. Facilitating the development of a post-disaster rehabilitation project for affected PRIME members, providing safe water and other supplies as available from government sources, and conducting cash-for-work activities to repair/maintain/develop community infrastructure
		<b>Love to see partner organizations</b>
		9. Receiving management training to establish a viable and RBM-driven branch office, and providing skills training to respective staff members to implement PRIME components



<b>Output</b>	<ul style="list-style-type: none"> <li>• Groups exist/are available as informal social institutions</li> <li>• Members access different formal financial services</li> <li>• Members use emergency loans to overcome short-term emergency crises.</li> <li>• Members who received training are capable of undertaking viable IGAs</li> <li>• Members utilize technical services after initiation of IGAs</li> <li>• Members or their family members are skilled in prospective informal trades (vocational training)</li> <li>• Access to primary healthcare services available to primary beneficiaries.</li> <li>• Ultra-poor from disaster-affected areas utilize rehabilitation services</li> </ul>	<p><b>Output indicators:</b></p> <ul style="list-style-type: none"> <li>• No. of households become group members (20–25 per branch), 90% of whom satisfy a set of quality criteria (regular meetings, issues-based discussion, etc.) within six months</li> <li>• 80% of members have received a loan three times, 60% of the loan is used for declared IGA purposes</li> <li>• All positively assessed members have received emergency loans to protect IGAs.</li> <li>• 60% of participants who received IGA training can explain, demonstrate and implement the technical process of an IGA</li> <li>• All members who requested technical services have received them</li> <li>• 80% of the trained participants are either employed or in internships</li> <li>• 80% of members can demonstrate awareness issues after sessions</li> <li>• 100% of patients of satellite clinics receive relevant basic healthcare services</li> <li>• 60% of referred clients have proof of having visited referral institutions</li> <li>• No. of disaster-affected households have access to potable water</li> <li>• No. of person-days for rehabilitation works (community infrastructure, maintenance work, etc.) created</li> </ul>
<b>Project activities</b> (project team)	<ul style="list-style-type: none"> <li>• Formation of groups of members</li> <li>• Facilitation of access to financial services</li> <li>• Enabling of access to emergency loans</li> <li>• IGA capacity building</li> <li>• Provision of technical services</li> <li>• Provision of vocational training</li> <li>• Provision of primary healthcare</li> <li>• Development of disaster management</li> </ul>	<p>Strategy map/journal (concept in OM for monitoring activity, an optional choice)</p>

There were strong capacity building and behavioural change elements in this project, which can be seen explicitly from some of the results statements in the above table, and they were also deeply embedded in the project operation and implementation process. Additionally, there was a clear structure and division of roles of different actors, such as easily identifiable boundary partners whose capacity and sustainability were critical for delivering planned services and achieving project goals. Moreover, changes in practices, attitudes and relationships among major actors at both individual and institutional levels were also anticipated results.

Two broad elements are accommodated in the first column of the matrix: 1) distinctive and explicit groups of actors – lead project team, boundary partners (defined on the basis of the lead project team as a reference point), and beneficiaries; and 2) a hierarchy of levels of results – activities, output, outcome, goal and impact –

each shown beside the actor to which it corresponds. The second column presents expected targets and results; and the third column accommodates performance indicators linked to each level of result with both quantitative and qualitative parameters.

As with Case 1, I referred to the original project logframe, monitoring reports, and longitudinal impact assessments to draw useful data when reconstructing the project framework with the synthesis model. In the original project logframe, there were three levels of results – output, outcome and impact. A distinctive feature of this logframe was that outputs were divided into three categories according to the project actor corresponding to each – PKSf-based outputs, partner organization-based outputs, and other target group (beneficiary) related outputs (PKSF, 2013). This shows that an actor dimension already existed in the original project logframe. However, when reformulating the project matrix with the synthesis model, I needed to resituate these three groups of project actors and redefine related results. Specifically, as shown in Table 6.2, I identified these partner organizations as boundary partners of the lead project team and then distinguished an additional level of result – outcome as behavioural and capacity change in partner organizations (one of the key objectives of the project but not sufficiently reflected and measured by the existing evaluation system, as mentioned in Chapter 4). In doing this, I also concluded the determination of my primary data sources – partner organizations – and other relevant project actors whom I interviewed during fieldwork.

Next, I extracted the data – particularly those on results and targets respectively linked to the three actor groups identified from secondary data – necessary to complete the project matrix for the synthesis model. This process was relatively more

straightforward at output, goal and impact levels as major relevant content could be transferred directly after necessary differentiation with reference to their connections to various actors. At outcome level, however, I needed to first gather and extract all the available information, including targets/results and indicators linked to boundary partners, and then attempt to abstract the main common objective to formulate the outcome challenge statement.

As mentioned earlier, there were partner organization-based outputs in the original project logframe and these data were integrated into the reformulated framework as boundary partner-related outcome and indicators. After my further categorization with reference to their different degrees of progression on the change/results ladder, these indicators formed the progress markers shown in Table 6.2. For example, the following outputs were listed in the original project logframe:

***Output 1: Group Formation***

- a) Select beneficiary following prescribed criteria
- b) Develop profile of selected beneficiaries
- c) Form groups to be guided by a committee
- d) Introduce flexible savings system
- e) Hold issues-based discussions in group meetings.

***Output 2: Access to financial services***

- a) Determine IGA-based credit needs
- b) Provide flexible credit facilities
- c) Follow up the use of loans and debt situations of members

***Output 3: Access to emergency loans***

- a) Assess emergency loan needs of IGA-loanees during the season of vulnerability
- b) Provide emergency loans (PKSF, 2013)

On the basis of this information and with reference to other project documents containing relevant information, I inferred and formulated the following progress markers for boundary partner groups shown in Table 6.2:

**‘Expect to see’ progress markers for boundary partner organizations:**

1. Selecting appropriate beneficiaries, developing profiles, forming groups, introducing a flexible saving system, and holding issues-based discussions in group meetings
2. Determining IGA-based credit needs, providing flexible credit facilities, and following up on the use of loans and debt situations of members
3. Assessing loan needs of IGA loanees during the season of vulnerability and providing emergency loans

These progress markers provided the basis for designing my interview guide and specific progress-marker-related questions (see Appendix 11). They proved to be relevant and were validated by respondents (especially members from partner organizations) through recalling changes during the interviews (relevant illustrative examples are given in later sections).

Compared with Case 1, the reformulation of the project framework in Case 2 was a relatively less time-consuming and challenging process. There were several reasons for this: 1) I had obtained some initial practical experience in the first case study, in addition to my theoretical knowledge and understandings prior to the empirical research; 2) there was a clear existing actor thread in the original project framework of Case 2, which proved useful in the process of identifying and differentiating actors under the synthesis model; and 3) there was a simpler actor structure in Case 2, as there was only one group of boundary partners (albeit of 24 organizations) whose member institutions came from the same micro-finance sector and had similar roles in the project (I return to this issue in Section 6.1.5).

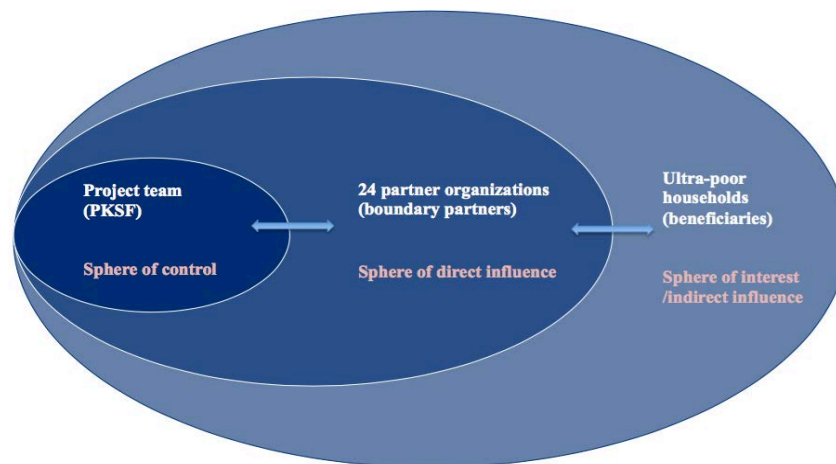
Following a similar case study analysis strategy and protocol as with Case 1, the following sections unpack the various components of the synthesis model evaluation framework for Case 2 in the sequence actor analysis, results chain analysis, and evaluation mechanism analysis.

### 6.1.3 Actor analysis

In accordance with the reformulated project framework for Case 2 in Table 6.2, I focus on the following three groups of actors in the actor analysis:

- Lead project team (PKSF)
- Boundary partners (24 partner organizations)
- Beneficiaries (ultra-poor households in targeted project areas)

As with Case 1, I took the lead project team in Case 2 – PKSF (working under direct supervision of the donor agency) – as the reference point in defining boundary partners, and adopted the same sphere of influence tool to develop the actor aspects of the theory of change of the synthesis model in Figure 6.1 below. This actor strand of the theory of change shows not only the chain of actors in each project domain, but also illustrates the sequence in which each was affected by and effected changes in others. PKSF (located in the sphere of control) worked directly with 24 partner organizations (boundary partners located in the sphere of direct influence) to reach the targeted ultimate beneficiaries (ultra-poor households located in a sphere further beyond the direct influence of the project team). In this project, PKSF supported boundary partners with operational subsidies, various types of technical assistance, and capacity building through management and staff training. Partner organizations were then expected to become self-sustaining in their own institutional operations, and capable of implementing the project in the field and delivering financial and non-financial services to targeted ultra-poor households more effectively.

**Figure 6.1: Actor strand of the theory of change in Case 2**

Moreover, relationships among these actors in different project domains were nonlinear and interactional under the synthesis model. For example, through various kinds of support strategies, supervision and monitoring mechanisms, PKSf exerted influence on boundary partners and effected changes in their behaviour, mindset and attitude during the project implementation process. Reaction and feedback from boundary partners in turn affected and informed the PKSf decision-making process with regard to adjustment of the programme implementation and monitoring strategy and the institutional development agenda. Finally, beneficiaries in this actor chain – targeted ultra-poor households – provided feedback on achievements and needs to frontline project staff during regular household visits and focus group discussions.

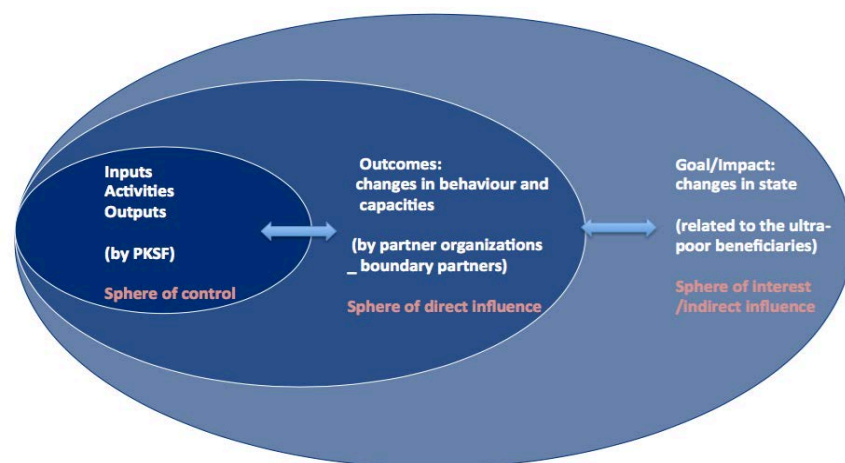
Utilizing this actor strand of the theory of change, which clarifies the actors involved, their respective project domains, and their interwoven relationships, I was able to identify individuals with whom to seek interviews with the aim of validating and deepening my understanding of how the project worked in practice. However, the actor component alone was far from sufficient to inform the evaluation design and further unpacking of the results chain and evaluation mechanism according to the

synthesis model was necessary.

#### 6.1.4 Results chain analysis

Following the above breakdown and analysis of the actor component under the synthesis model, I sought to explore various levels of results in connection with these actors to reveal changes which had occurred. Using the sphere of influence diagram as an analysis tool, I sought to visualize the boundaries between the results domains in the synthesis model, and to build the results strand of the theory of change in the form of a results chain, as shown in Figure 6.2 below.

**Figure 6.2: Results strand of the theory of change in Case 2**



As I applied the synthesis model to two case study projects, the results chain exemplified in the above figure includes elements similar to those of Case 1: project outputs, outcomes, goal and impact-level results (see detailed content in Table 6.2) located in different project domains respectively. As mentioned earlier, capacity building was a key element in Case 2, both as a means – the process boundary partners needed to go through to gain the necessary capacity – and as ends – the intermediate outcomes targeted and generated by the project. In the above results chain, outputs (related to PKSf), goal and impact-level results (changes at ultra-poor

household level) are associated with more traditional dimensions similar to those of the LFA or RBM; while outcome-level results are defined as behavioural and capacity changes in boundary partners.

Moreover, in the results chain formulated under the synthesis model, results in relation to beneficiaries are stratified in terms of two levels: the goal and a more transformative level of change state – impact. This illustrates the embedded critical realist perspective in evaluation under the synthesis model in which stratified reality is represented by layered and gradually progressive results. In addition, as the gradually fading shading in each project domain in Figure 6.2 indicates, the influence of the project decreases along the chain of results and actors, and the owners of change shift from the lead project team (PKSF) to intermediary boundary partners and local ultra-poor households.

#### **6.1.5 Evaluation mechanism analysis**

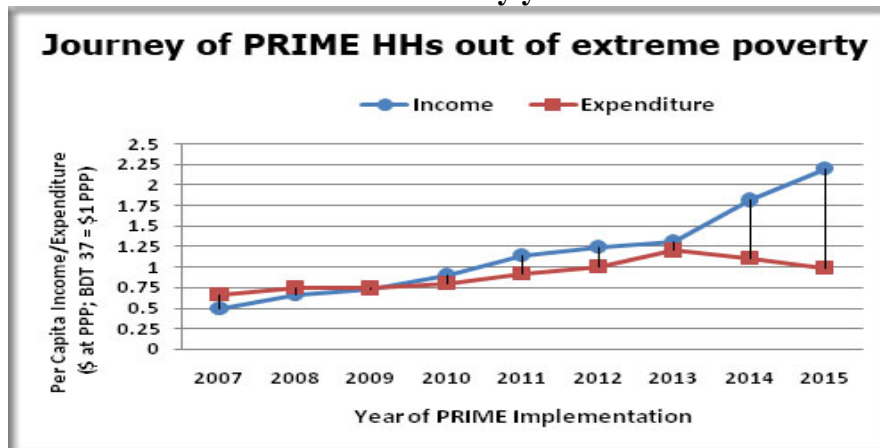
Utilizing the actor and results chain analysis, I sought to unpack the process by which the synthesis model could be applied in evaluating the second case study project and the major components which needed to be identified and defined first. Accordingly, in this section, I consider how the evaluation mechanism of the synthesis model could be used in this case with some illustrative empirical examples. Here, I am using the term ‘evaluation mechanism analysis’ in the same way as defined in Chapter 5: the process and means involved in guiding and operationalizing evaluation data collection and interpretation. This specifically involves the instruments accommodated in the synthesis model for evaluation and the various options and possibilities for analysing and making sense of the data collected.



As discussed in the first case study analysis in Chapter 5, the evaluation mechanism under the synthesis model is operationalized mainly through packages of performance indicators related to each level of results and the corresponding actors. There are output indicators, outcome challenge/progress markers, goal indicators, and impact indicators. I referred to the indicators and changes to be examined in the reformulated project matrix in Table 6.2 when collecting primary evaluation data for Case 2. Outputs, goal and impact-level results were mainly gathered from secondary data comprising quantitative and some qualitative information (examples showing goal and impact-level results are shown in figures 6.3 and 6.4), supplemented and verified by some of the primary data gathered through my interviews with the lead project team and staff from the donor agency.

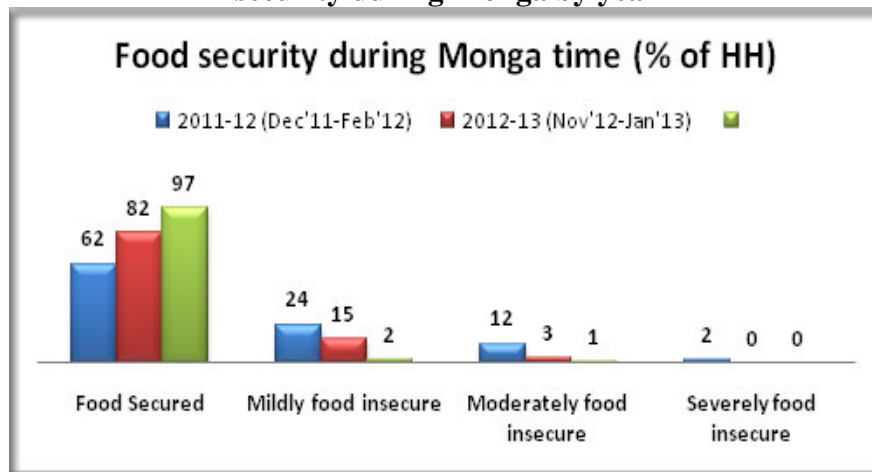
Outcome-level results were collected through (mainly qualitative) primary data with reference to the outcome challenge and progress markers, as specified in detail in Chapter 4. In this project, I only defined one set of progress markers, as there was only a single group of boundary partners comprising 24 partner organizations. Through the tracing of behavioural change in these boundary partners (from their recall of change stories in interviews), I then sought to use this primary data as another important piece of evidence to tell the performance stories, supplement other levels of results, and fill the gap between outputs and goal.

**Figure 6.3: Changes in income and expenditure of ultra-poor households by year**



(Note: HHs = households)  
(Source: [PKSF, 2013; 2017](#))

**Figure 6.4: Project households at different levels of food security during Monga by year**



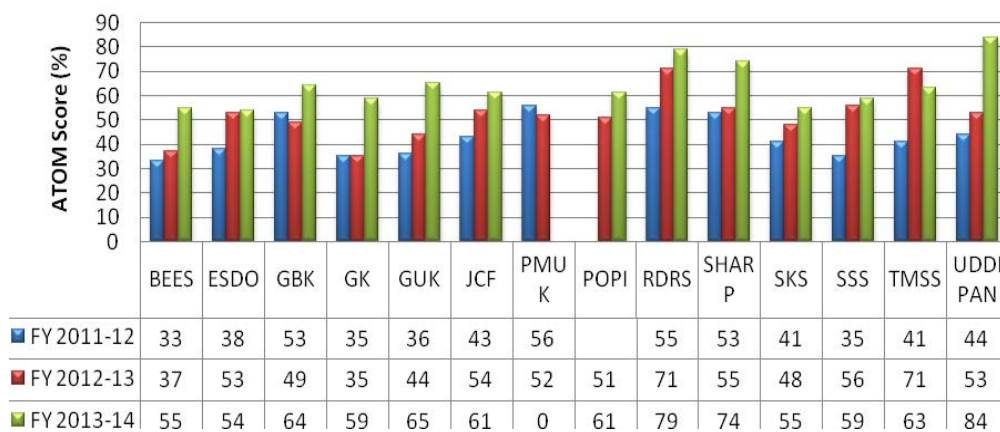
(Note: HHs = households)  
(Source: [PKSF, 2013; 2017](#))

As indicated in the previous chapter, there are different ways of analysing and making sense of the data gathered through the synthesis model to form evidence chains and to serve specific evaluation purposes. The evaluation data analysis and interpretation principles mentioned in Chapter 5 can also be applied to Case 2, such as analysis of the final results with reference to a baseline parameter or situation, or analysis from a grounded theory perspective.

For example, as discussed in Chapter 4 and shown in the summary of key features of Case 2 in Table 6.1, there was an established RBM system in place as part of the internal evaluation system in this project. This system utilized assessment tools such as activity to output monitoring (ATOM<sup>38</sup>) and the Economic Self-sufficiency Assessment Score (ESSAS<sup>39</sup>). According to collected secondary data from an RBM report, the project mainly studied and assessed partner organization-level achievement and performance with these two tools and the resultant ATOM score and ESSAS of 14 partner organizations<sup>40</sup> are respectively indicated in figures 6.5 and 6.6. Using scores from these means of assessment, the project then identified four under-performing partner organizations (PKSF, 2013).

**Figure 6.5: Partner organization ATOM score**

**PO-wise Activity to Output Monitoring (ATOM) Score**  
(EMS = 69.8; Error df = 15;  $P < 0.01$ )



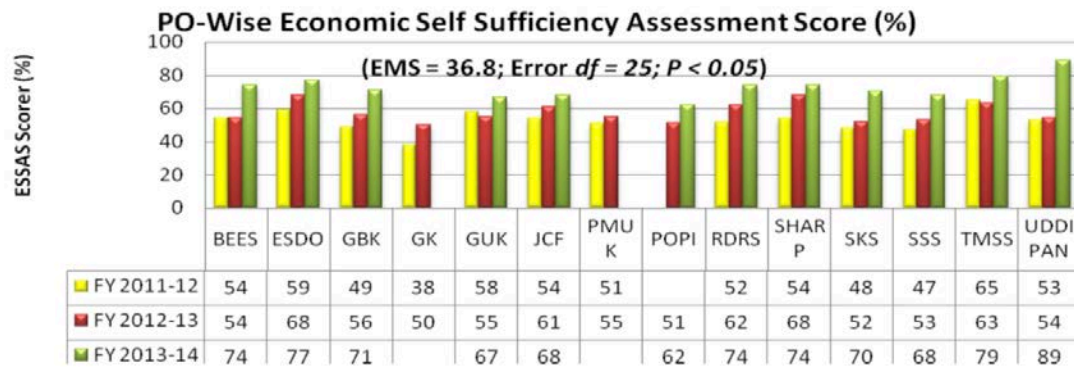
**Note:** PO = partner organization

(Source: PKSF, 2013)

<sup>38</sup> At output level, ATOM system uses time-based, cumulative percentage progress towards overall targets for each element of the programme (PKSF, 2013).

<sup>39</sup> At outcome level, ESSAS is used to measure the economic self-sufficiency of PRIME beneficiaries that will help them fight food insecurity during Monga periods (PKSF, 2013).

<sup>40</sup> Partner organizations are represented by abbreviations along the x-axis of each graph.

**Figure 6.6: Partner organization ESSAS**

**Note:** PO = partner organization

(Source: PKSF, 2013)

However, the rationale behind such a relatively low performance and the pathway of change remain unclear. This missing thread of the story could be captured by the actor-oriented synthesis model, under which I identified these partner organizations as boundary partners and gathered behavioural and capacity change stories from them together with possible influencing factors (as discussed in Chapter 4).

There follow extracts from interviews with project officers from these under-performing partner organizations on influencing factors and challenges that might have affected project implementation:

Some beneficiaries are non-educated, so it's difficult to communicate with them [about project concepts] and they did not have the motivation for participation at the initial stage; as project staff, we need to make a lot of effort to motivate them. And most of the people are landless, which is also a difficulty in implementing agricultural-based income-generating activities. Another thing is that they lack skills, and because they are not educated, they forget the skill development training we provided in one year; after one year, we need to train them again. It's difficult to sustain. (Boundary partner interview 9, May 2015, Rangpur, Bangladesh)

We have very little organizational challenge at PO level. But some of our beneficiaries live in remote areas, and they live here and there, it's difficult to reach and gather them. There is migration issue: some people borrow money and then they move from one place to another, difficult to track and collect loan

repayment from them. Because of the low capacity of poor people, there is sustainability issue: they have no capacity in using technical and micro-credit services provided to them, and it's difficult to sustain the outcome of training... We faced problems caused by natural climate, flood, drought, and also political unrest. For example, when strike happens, it stops staff going to field. And at flood period, it's difficult for staff to commute to the beneficiary site, and people lose some products from income-generating activities, and disease happens at this time. (Boundary partner interview 32, June 2015, Dhaka, Bangladesh)

Figure 6.7 shows a word cloud I generated by using NVivo software on qualitative interview data on the difficulties and challenges these partner organizations encountered when implementing the project. The words displayed in the cloud indicate the most frequently mentioned key areas of challenge and difficulty. For example, *dropout problems* signify a high beneficiary and staff absence or resignation rate; *migration* shows high beneficiary mobility, which increased the difficulty in collecting loan repayments; *education* implies difficulty in communication and understanding due to the education level of beneficiaries.

This word cloud was generated mainly as an illustration of some interview data. These aforementioned practical problems helped indicate the rationale behind interviewees' low performance to some extent, but such challenges did not only apply to them. Triangulation of various sources of information was necessary to build the storyline and evidence chain by combining actual behavioural changes, internal institutional operation, and specific climatic conditions in relation to geographical location (e.g. particular seasonal changes in some regions of Bangladesh, as discussed in Chapter 4). In this way, better-informed decisions could be made on the adjustment of supportive strategies and performance improvement plans for partner organizations.

This example shows the added value of the inclusion of an additional level of results

on the behavioural changes in partner organizations (especially in combination with the exploration of influencing factors) – unpacking the change process and explaining the rationale behind the performance measured quantitatively.

**Figure 6.7: Frequently mentioned key words in partner organization interviews**



As mentioned in Chapter 5, the primary focus of this study was not to look at how a project performed and then produce a comprehensive evaluation report, but to draw implications on the applicability, utility, value added, and limitations of the synthesis model through empirically applying it in two case studies. These illustrative examples drawn from Case 2 reveal that the synthesis model can accommodate both quantitative and qualitative means of collecting and making sense of data, and help supplement and triangulate them to formulate comprehensive storylines and evidence chains.

## **6.2 Uncovering the distinctive value in applying the synthesis model**

Following the unpacking and analyses of the major components of Case 2 under the

evaluation framework of the synthesis model, in this section, I examine three major aspects with the same rationale and aim as specified for Case 1 in Chapter 5: an extended discussion of the actor element, linking and comparing the evaluation system of the synthesis model with those of the LFA and other evaluation approaches originally used in the Case 2 project, and a discussion of the influencing factors affecting various levels of results under the synthesis model in this project.

### **6.2.1 An extended discussion of the actor component in Case 2**

I applied the same actor structure of the synthesis model to both case study projects: three main categories – lead implementation team, boundary partners, and beneficiaries. The lists of specific actors varied in accordance with the different type and objectives of each project respectively.

In Chapter 5, the possible ‘messy partnership’ (Guijt, 2008) among various groups of actors was discussed in respect of Case 1. As there were five member institutions in the lead project team and four groups of boundary partners from different sectors in the first case, it was more likely that the partnership between them would be relatively complex due to considerable diversity in governance structure, capacity, level of commitment, etc. In Case 2, although the project scale was much larger, the composition of actor groups was relatively simple: only one institution in the lead project team, and a single group of boundary partners comprising 24 partner organizations from the same micro-finance sector with similar mandates, and organizational and governance structures. As these boundary partners also implemented similar packages of project activities aiming at achieving common objectives, there might be less complexity in the partnership between them. However, a certain level of messiness may exist within the same group of actors – such as these

24 partner organizations in the boundary partner group – given that their varied capacities, skills and commitment to the project vision would more or less affect the outcomes of the partnership, and the fact that they were working in different regions of the country where different contextual factors prevailed.

Furthermore, in Case 1, according to original project documentation, the actor component was relatively abstract under a single results-chain-focused theory of change. On the other hand, Case 2 had an actor thread in its original logframe which was mainly linked to output-level results. Under the synthesis model, I reconstructed this actor element by providing more explicit division and linking it to all four levels of results. This procedure provides a clear indication of change/task owners and results domains in evaluation, which then enables either a comprehensive or a tailored evaluation design in practice.

Finally, at an operational level, the actor structure determines the evaluation design and data collection process. For example, in Case 2, at the qualitative interview stage, I only designed one set of questions to interview all respondents from partner organizations, as they were members of the same boundary partner group (see Appendix 11). Conversely, four sets of questions were needed for interviewing boundary partner respondents in Case 1 (see appendices 5–8).

Consideration of the actor component and related issues in both cases indicates that the synthesis model can be applied flexibly to projects with largely different actor compositions. However, this is not to generalize the usage of the model on the basis of two case studies: each instance should be considered separately in case-specific situations, and it is critical to draw boundaries carefully and clearly differentiate roles



when defining various groups of actors.

### **6.2.2 Linking and comparing the synthesis model with the LFA and the existing evaluation framework**

As indicated in Table 6.1, the LFA was originally adopted in Case 2 not only in project planning but also as the central component of its RBM system in performing internal self-evaluation together with other instruments. This can be seen from the following extract from a project monitoring report:

PRIME introduced Results Based Monitoring (RBM) from FY2010–11... In the RBM system, the principle focus is on results – output, outcome and impact – using the logical framework as the ‘heart’ of the system. At output level, the activity to output monitoring (ATOM) system uses time-based, cumulative percentage progress towards the overall target for each element of the programme. At outcome (purpose) level, Economic Self-Sufficiency Assessment Score (ESSAS) is used to measure the ‘economic self-sufficiency’ of PRIME beneficiaries that will help them fight food insecurity during Monga periods. At Impact level, modified USAID Household Food Insecurity Access Scale (HFIAS) combined with Monga coping strategies provide the seasonal vulnerability reduction index (SVRI) of PRIME HHs [households]. (PKSF, 2013)

Nevertheless, this RBM system including its LFA element and other tools mainly employed numerical means and gathered quantitative evaluation results which did not make explicit how and why changes occurred. The synthesis model aimed to provide these explanations by examining micro-level behavioural and capacity changes in intermediate actors.

In addition to the internal self-evaluation system, there were also other forms of evaluation required by the donor agency. For example, a series of longitudinal impact studies were undertaken by other external and independent institutions to assess 1) beneficiary-level results through a quasi-experimental evaluation design; and 2) partner organization-level results through a sustainability analysis defined by

operational self-sufficiency (OSS) and financial self-sufficiency (FSS)<sup>41</sup> ratio as exemplified in Table 6.3:

**Table 6.3: Self-sufficiency ratios of partner organization branches by year**

Year	Operational Self-sufficiency				Financial Self-sustainability			
	Mean	Median	Standard Deviation	% Branches Sustainable	Mean	Median	Standard Deviation	% Branches Sustainable
2008	0.7986	0.7253	0.4629	26.80	0.6131	0.5743	0.3479	13.07
2009	0.8839	0.7910	0.5120	30.29	0.7023	0.6522	0.3498	15.92
2010	0.8298	0.7257	0.4345	29.38	0.6963	0.6464	0.3389	15.12
2011	0.9545	0.9072	0.3620	35.98	0.7634	0.7358	0.2700	13.53
2012	1.0018	0.9978	0.3986	49.54	0.8119	0.8099	0.2843	22.27

(Source: Khalily et al., 2013)

At the time of the study, the latest round of the longitudinal impact assessment had updated its approach to include the additional dimension of human dignity on the basis of early-stage parameters such as food security, economic wellbeing, a multidimensional poverty index, etc. This human dignity and respect dimension was measured in terms of five parameters: (i) social status and empowerment; (ii) household decision making; (iii) women's status in the community; (iv) overall awareness; and (v) respect and dignity (Khalily et al., 2016).

The human dignity and respect dimension added by the external impact study team was certainly valuable in exploring and measuring other significant aspects of the project. However, it still focused primarily on the beneficiary level with mainly quantitative means of measurement and did not directly address the need to assess the capacity building element particularly at institutional level (e.g. in relation to partner organizations), which was another key objective of the project, as mentioned in Chapter 4. The evaluation mechanism under the synthesis model served as a feasible

<sup>41</sup> OSS is defined as the ratio of total operating revenue (excluding direct and cash subsidies) to total expenses (including the loan loss ratio). FSS ratio is defined as the ratio of total operating revenue to adjusted total expenses (Khalily et al., 2013).

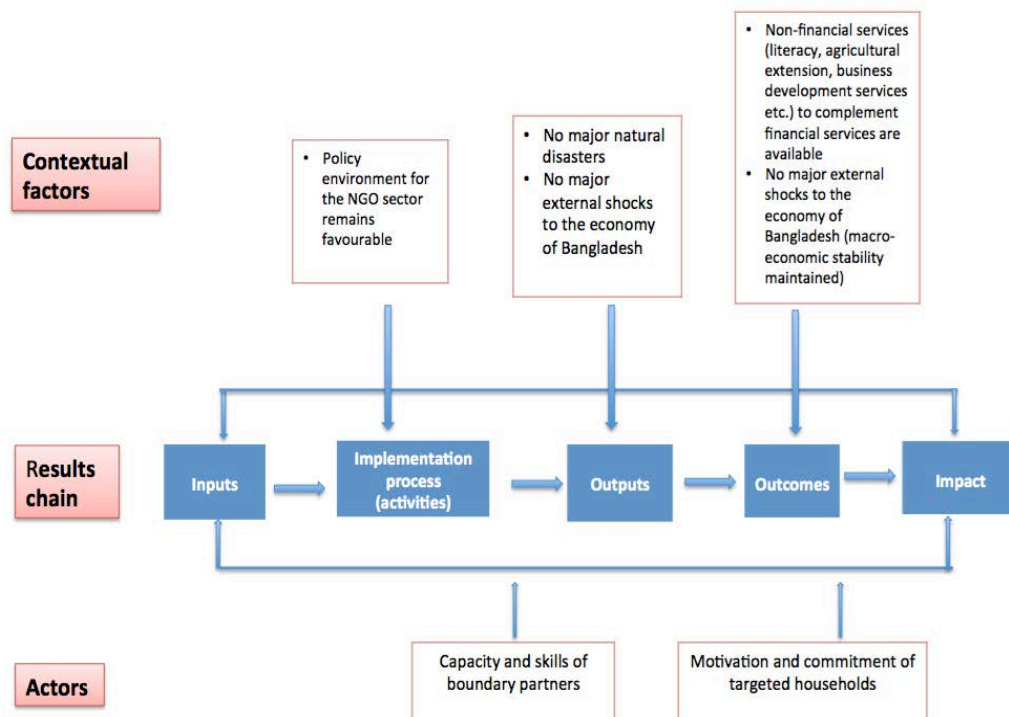
means of filling such a gap and the data thus obtained could well supplement those from other studies (see examples in Section 6.1.5 of this chapter).

### 6.2.3 Analysis of influencing contextual factors in Case 2

This section discusses the influencing and contextual factors at play in Case 2 to explore their relationships with various levels of results and to show the case-specific conditions under which the synthesis model was applied.

By exploring and extracting contextual information from the assumption column of the original project logframe and other project documents, I constructed Figure 6.8 to indicate the major influencing factors in connection with different activities and results along the results chain under the synthesis model.

**Figure 6.8: Analysis of influencing factors in Case 2**



For example, the policy environment for the NGO sector had a critical influence on

the implementation of project activities and was one of the major concerns that project stakeholders identified during the project inception stage. The following extract is from an early-stage project document which analysed the policy environment for NGOs and the potential risks involved:

Bangladesh presents some unique challenges to those working in any area of development, whether they are an NGO, civil society group, or individual. Reaction to perceived attacks or partisan activities can be quite swift and in some cases severe. There has also been some evidence of a shift in Government policy towards new NGOs. Information obtained from officers managing local donor funds suggests that the registration of new NGOS has become quite difficult, with some substantial time delays. This is reportedly the result of a sub-rosa directive from the Government who wish to control growth of this sector. (DFID, 2006)

In addition, environmental factors such as a major natural disaster could directly affect outputs related to emergency loan disbursement or the provision of rehabilitation services to the ultra-poor. Another contextual factor – a major external economic shock to the country – might influence multiple levels of results including outputs and outcomes. Domestic economic stability is a key factor in fostering a sustainable micro-finance sector, which then directly impacts on partner organizations (micro-finance institutions) in terms of sustainability as well as service delivery to beneficiaries. Although some contextual factors have more direct influence on one level of result than another, they are often interwoven and impact on the overall results chain.

I also distinguished the actor-specific factors under the synthesis model: the capacity and skills of boundary partners, and the level of motivation and commitment of targeted beneficiaries also have an impact on relevant levels of results and project performance as a whole. An actor-specific analysis of influencing factors is feasible with the prior differentiation of actors under the synthesis model.

The analyses of influencing factors in both case studies not only facilitate in-depth understanding of specific conditions and risks the project may have encountered in the implementation process, but also help indicate the applicability of the synthesis model in diverse contexts so that ‘under-similar-conditions’ generalization can be generated as a reference for future application of the model.

In both case studies, the systematic unpacking of the synthesis model application process in terms of reconstructing the overall project framework and the breakdown of its various components – actors, results chain, and evaluation instruments – revealed the model’s evaluation function. The model provided a framework for constructing the double strands of the theory of change – the actor strand and results chain – which clarified two key aspects of evaluation, namely, who (the major actors focused on) and what (the specific results/changes examined). These two components together with corresponding indicators included in the project framework formed the instruments to operationalize the evaluation of each project in practice. The value added, and limitations and implications arising from the practical application of the synthesis model are summarized and reflected on in the final chapter to address the research questions and draw some conclusions.

## Chapter 7 Conclusions

In this chapter, I summarize and conclude with answers to the overarching research question on the evaluative insights generated in constructing and applying the synthesis model with reference to three sub-questions as set out in Chapter 4:

- 1) How does the synthesis model, as an alternative evaluation approach, respond to identified evaluation challenges and demands – in particular how well does it reconcile the tension between learning and accountability, offer explanatory power about the links between outputs and outcomes or goal and add an actor perspective?
- 2) What value is added and what limitations arise in adopting the synthesis model in practice?
- 3) What are the lessons and implications for further research and evaluation practice?

To answer these questions, this chapter includes 1) a general appraisal of the synthesis model as a theory-based evaluation framework with a mixed methods design; 2) responses to certain contemporary evaluation challenges and demands, as featured in the conceptual framework; 3) the practical value added by and limitations of the synthesis model; and 4) reflection on the synthesis model in terms of its application, lessons learnt, and implications for further research and evaluation practice.

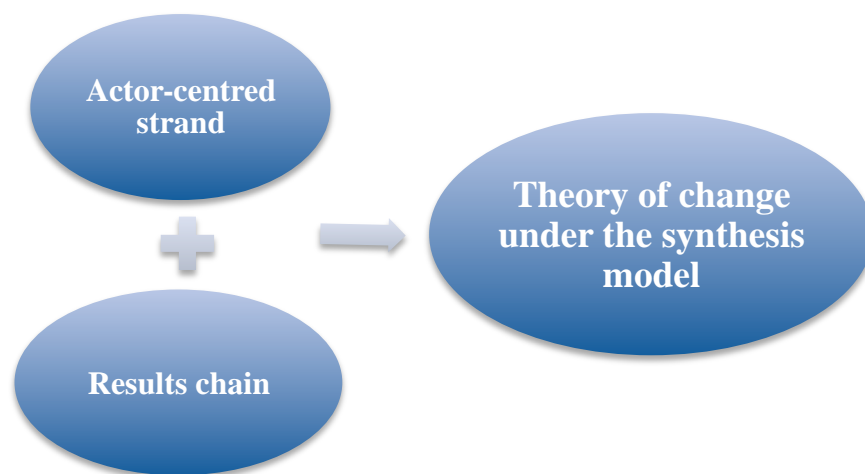
### **7.1 The synthesis model – a theory-based, mixed-methods evaluation framework**

In Chapter 3, I discussed the potential of the synthesis model to provide a holistic theory-based evaluation framework with a double-stranded (actor strand and results chain) theory of change, which enables either a comprehensive or a tailored evaluation. The model also provides a systematic framework for the breakdown of a

development project into various components, including ladders of change/results with their corresponding indicators and groups of actors. It thus accommodates a mixed methods evaluation design that utilizes specific tools to assess different components.

These theoretical assumptions were examined empirically in my application of the synthesis model in two case study projects. As discussed in chapters 5 and 6, in both cases, the model proved capable of serving as a theory-based evaluation framework functioning through the double-stranded theory of change system, as shown in Figure 7.1 below. By following the logic of the synthesis model, the chain of actors identified in both case studies (project team, boundary partners and beneficiaries with their respective roles and responsibilities) forms the first strand of the theory of change and indicates the major actors to be focused on in evaluation; while the second strand – in the form of a results chain – further identifies the types of changes/results to be evaluated with their interrelated relationships along the chain.

**Figure 7.1: Double-stranded theory of change under the synthesis model**



(Source: The author)

These two aspects of the theory of change function together to guide the evaluation design and data collection strategy, such as a mixed methods design for a relatively complex case. This can be seen from the application of the synthesis model in both case studies. For example, in Case 1, I gathered output-level results through LFA instruments (mainly quantitative, pre-defined targets and indicators); outcome-level results against qualitative indicators – progress markers – and qualitative data collection methods such as interviews; and impact-level changes against quantitative parameters. Similarly, in Case 2, I employed qualitative methods for outcome-level evaluation to gather data on behavioural and capacity changes in intermediary actors, which supplemented impact-level assessment conducted by the project through a quasi-experimental design. This shows that the synthesis model is suited to a mixed methods evaluation design on the basis of a well-articulated theory of change.

Furthermore, with the holistic evaluation framework available under the synthesis model, there is certain flexibility in selecting either a comprehensive or a tailored design according to specific evaluation purpose, priority and questions. This too proved feasible in practice when dealing with different evaluation needs and focuses at different stages in each case study project. For example, in Case 2, as mentioned in previous chapters, there was a results-based monitoring system in place throughout the implementation stage, which reflected a tailored design to assess outputs in relation to project activities and the results on economic self-sufficiency of project beneficiaries; there were also a tailored outcome-level evaluation particularly focusing on capacity and behavioural changes in intermediary actors and independent impact-level longitudinal studies. All these separate evaluations at various project stages with different focuses can be embedded in the overall framework of the synthesis model.



## **7.2 Responses of the synthesis model to identified evaluation challenges**

In this section, I seek to answer research question 1 in terms of the responses of the synthesis model to 1) reconcile the tension between learning and accountability; 2) add an actor perspective; 3) offer explanatory power through unpacking mechanisms or links connecting outputs and outcomes or goal in evaluation.

### **7.2.1 Reconciling the tension between learning and accountability**

As discussed in the literature review in Chapter 2 and as reflected in the conceptual framework, the tension between two goals of evaluation – learning and accountability – and the difficulty in achieving both concurrently remains a key challenge. Indeed, with more emphasis on learning for improved development practice in addition to the traditional central focus on accountability, an appropriate evaluation approach or framework which can accommodate and moderate both goals becomes particularly necessary.

The synthesis model can resolve this issue through combining the results orientation of the LFA – which responds to the need to provide financial accountability to donor agencies – and the process pathways of OM – which address the need to generate and support learning for programme actors and donors (the basic assumption of the original authors in designing the synthesis model). As discussed in Chapter 3, evaluation through the LFA represents a goal-oriented approach which has the necessary structure and processes in place to achieve mainly upward, financial, accountability. Therefore, the synthesis model can rely on its LFA elements to help answer questions on whether certain goals are achieved against set targets and whether services are delivered accordingly. Moreover, the quantitative indicators of

the LFA can be more easily aggregated than those of OM to satisfy simple reporting needs. The synthesis model is also capable of providing an instrument – a comprehensive double-stranded theory of change – to help achieve internal and downward accountability through accommodating and focusing on a wider range of actors including implementing organizations and beneficiaries, as well as related results. The potential of the synthesis model in realizing internal and downward accountability was not fully explored empirically as it needs longer-term and more regular involvement with the project and its various participants including beneficiaries, which is beyond the scope of this PhD study. On the other hand, with the process-oriented learning pathways reflected by its OM components, such as outcome-level results (defined as behavioural change in boundary partners), outcome challenges, and progress markers, the synthesis model also enables learning.

The application of the synthesis model in both case studies reflected this analysis. First, in terms of accountability (mainly upwards), in Case 1, in addition to the mechanisms of reporting on financial aspects (e.g. budget reports), output-level evaluation (e.g. assessment of adherence to donor requirements) adopted largely accountability-based criteria such as efficiency and effectiveness. In my application of the synthesis model as detailed in chapters 5 and 6, I defined output-level results in terms of a similar notion to that of conventional LFA usage, operationalizing evaluation through preset targets and indicators. Likewise, the model enabled accountability-based evaluation in Case 2 with similar LFA-related elements, particularly at output level in the results chain.

Second, in terms of learning, following the results chain in the synthesis model, I defined outcome-level results as behavioural change in boundary partners, which

helped in unpacking the process of how micro-level results – outputs and outcomes – contributed to macro-level change – goal and impact – thus indicating the process-oriented learning pathway added by OM components. For example, as shown by the breakdown of outcome-level results for Case 2 in Table 7.1, these various elements helped fill the missing gap between outputs (in relation to the project team) and the project goal or impact (related to beneficiaries) with specific focus on the intermediary process. Thus, through the tracing of changes in behaviour, practice and capacity in relevant actors, the implementation process and generative mechanisms at play were revealed. Tracking progress markers would also involve continuous participation and reflection on the part of key project and boundary partners (particularly if OM elements were adopted from the project planning stage), which could help create a learning space. All these mechanisms can operate in harmony under the synthesis model and contribute to the achievement of the learning objective of evaluation.

**Table 7.1: Example of the contribution of OM elements to the learning pathway**

<b>Project outcome</b> (behavioural change in boundary partners)	<b>Boundary partners</b> (partner organizations)	<b>Outcome challenge:</b> Through various operational and capacity building support from PKSf, partner organizations are expected to become not only self-sustaining, but also capable of delivering both financial and non-financial services to their clients, including ultra-poor households, even beyond the project time span.	<b>Progress markers</b>
			<b>Expect to see partner organizations</b>
			1. Selecting appropriate beneficiaries, developing profiles of them, forming groups, introducing a flexible saving system, and holding issues-based discussions in group meetings
			2. Determining IGA-based credit needs, providing flexible credit facilities, and following up on the use of loans and debt situations of members
			3. Assessing loan needs of IGA loanees during the season of vulnerability and providing emergency loans
			<b>Like to see partner organizations</b>
			4. Identifying training needs for viable IGAs of selected members, providing practice-oriented IGA training (skills, market linkage), and establishing linkage between members and markets
			5. Assessing technical service needs of IGA practitioners and providing/arranging necessary vaccinations
			6. Preselecting suitable trainees for selected informal vocational training and following up on provision of training
			7. Undertaking door-to-door primary health care services, delivering selected medicines, providing referral to specialized and high-level government organization, NGO and private facilities

			8. Facilitating the development of a post-disaster rehabilitation project for affected PRIME members, providing safe water and other supplies as available from government sources, and conducting cash-for-work activities to repair/maintain/develop community infrastructure
			<b>Love to see partner organizations</b>
			9. Receiving management training to establish a viable and RBM-driven branch office, and providing skills training to respective staff members to implement PRIME components

In both case studies, the goal of learning from evaluation was fulfilled through engaging and interviewing key project actors such as boundary partners. As discussed in chapters 5 and 6, I not only gathered stories of boundary partners' achievements towards progress markers but also their rationale for effective or inadequate performance, in order to inform decisions on project adaptation and improvement.

Other components such as organizational practice and strategy maps included in the synthesis model (as briefly mentioned in Chapter 4) can also be utilized for internal accountability and organizational learning purposes. They are best explored and adopted from the project design and planning stage, and necessitate continuous involvement and collaboration among the various groups of actors identified by the model.

The synthesis model thus proved to be capable of combining and delivering well in terms of both (upwards financial) accountability and learning in evaluation by retaining the LFA function to fulfil the upward accountability objective and bureaucratic requirements of donor agencies, and integrating some key OM components to help unpack and explain the change process to achieve learning. Thus, under the synthesis model, the two goals of evaluation are harmonized in a mutually supportive fashion.

### **7.2.2 An actor perspective in the synthesis model**

As mentioned in Section 7.1, the synthesis model explicitly incorporates an actor-centred strand in the theory of change, which helps in identifying various groups of actors involved and their respective project domains, and differentiating their different roles and responsibilities. The synthesis model, in comparison to the LFA or OM alone, fills the actor gap and balances the actor focus – it explicitly brings a layer of intermediary actors into the chain, and considers the contribution of their behaviour and capacity dynamics. This makes the actor strand of the theory of change more comprehensive with a balanced focus on various groups of actors including project team, intermediary actors and beneficiaries. This actor-centred aspect supplements the results chain (as exemplified in the LFA framework) and the two jointly form a systematic theory of change.

As discussed in chapters 5 and 6, both case studies well utilized this actor-centred approach during the application of the synthesis model – actor aspects of the project theory of change were articulated in both cases to help identify and clarify relevant actors, their project domains and interactional relationships. In doing so, I got to refine the focus of the evaluation and distinguish various sources of different categories of evaluation data – whose perspectives on change I should gather and analyse in practice.

Moreover, the extended discussion of the actor component in both cases in chapters 5 and 6 indicate further benefits of an embedded actor perspective in the synthesis model, some of which correspond to the literature reviewed in Chapter 2 in this respect. For example, under the synthesis model, clarity on the actor element enables recognition of diverse characteristics and perspectives of different groups of actors

and thus affords better possibility to have in-depth understanding of their needs, interactions and relationships. This in turn brings better opportunities to achieve multiple levels of accountability – with various accountable actors and their responsibilities and relationships identified – and learning through tailored and focused support strategies and activities, especially in multi-actor contexts such as the two case study projects.

Additionally, from an actor-oriented perspective, a further advantage of the synthesis model is that it enables an actor-specific analysis of various influencing factors, as illustrated by figures 5.4 and 6.8. This has implications for tailored project strategies based on better understanding of the critical conditions that may affect project implementation and performance.

### **7.2.3 Explanatory power of the synthesis model – unpacking mechanisms linking outputs and outcomes or goal in evaluation**

As mentioned in Chapter 3, the synthesis model has the potential for offering or strengthening the explanatory power of evaluation through unpacking the intermediary mechanisms connecting outputs and outcomes or goal.

In practice, the application of the synthesis model in evaluating Case 1 brought an additional approach and perspective in reviewing project results, and helped fill the gap between outputs and goal through examination of the change process from an actor-centred perspective. In Case 2, there was originally no mechanism for reviewing the key capacity building element of the project particularly in relation to intermediary actors. The synthesis model helped bridge this gap through examining outcome-level results defined as behavioural and capacity change in boundary partners. In both cases, the model contributed to the building of an evidence chain

from minor behavioural change to macro-level transformation in corresponding groups of actors. It also made explicit three key aspects of evaluation: change and performance stories (what changed), change pathways (the process of how changes occurred), and the rationale behind change (why changes occurred).

Thus, the application of the synthesis model in both cases suggests that it is capable of offering explanatory power through unpacking the intermediary mechanisms and links (in the form of outcome-level results) connecting outputs and goal and balancing the results chain. By doing so, it meantime puts a mechanism in place to address process issues from an actor-centred perspective – enabling a process evaluation if required, which remains challenging in evaluation with the LFA alone.

Additionally, as reflected in the reformulated project frameworks of both cases in chapters 5 and 6, the synthesis model recognizes and accommodates the complexity of the development process through its holistic design with a relatively comprehensive package of both quantitative (e.g. as obtained through the LFA) and qualitative indicators (e.g. in the form of gradual progress markers for each intermediary project actor group). Other indicators measuring service delivery and organizational performance can also be incorporated into the evaluation framework of the model if necessary. This comprehensive system of performance indicators helps provide essential instruments for reflecting and assessing above-mentioned mechanisms and links between outputs and goal and thus helps gain strong explanatory power at an operational level.

### **7.3 Additional value added and limitations arising in applying the synthesis model in practice**

In this section, I address research question 2 in terms of two aspects: the additional value added, and limitations and challenges arising during the practical application of the synthesis model.

#### **7.3.1 Value added by the synthesis model**

As discussed in Chapter 2, irrespective of the primary focus or goal of the evaluation, attribution remains one of the basic practical challenges, and has become more challenging under the current results-focused development agenda. It was also noted that attribution may pose different challenges at different levels of results (e.g. outputs, outcomes, goal and impact).

It was also pointed out in Chapter 2 that the attribution issue is connected to the key elements of the conceptual framework – an explicit actor perspective and appropriate evaluation approaches with a degree of explanatory power can help address the issue to a certain extent, and tackling the issue has implications for achieving accountability through evaluation.

First, attribution at the output level – establishing logical and causal links between activities and outputs – is regarded as relatively straightforward (as discussed in Chapter 2) and can largely be addressed through the LFA-like attributes of the synthesis model. As reviewed in Chapter 3, the LFA not only helps identify expected causal links in the results chain, but also provides technical parameters (performance indicators) against which various project stakeholders can reach agreement in order to evaluate outputs and outcomes.



Application of the synthesis model in Case 1 reflected this relative ease in claiming and reporting attribution at output level with the LFA component. As citations from project documents in Section 5.2.2 revealed, logframe indicators were effectively used as references and parameters for output-level reporting, as required by the donor agency. Official responses from the project team showed clearly and directly established links between activities and results, as exemplified in Chapter 5. Output-level attribution in Case 2 was similarly straightforward: as indicated in chapters 4 and 6, there was an RBM system in place in which the LFA was the central component and instruments such as an activity-to-output monitoring score were adopted by the project to establish relevant logical links.

Second, attribution at outcome and goal or impact level – building causal links between the intervention and outcomes or impact – is considered to be challenging in development evaluation, as noted in Chapter 2. Some approaches have been developed to address this such as Contribution Analysis and OM (as discussed in Chapter 3). Both adopt the notion of contribution as an alternative to claiming attribution particularly in development impacts.

In adopting the OM definition of outcome and elements such as boundary partners and progress markers, the synthesis model addresses attribution at outcome level by assessing the contribution of boundary partners' behavioural and capacity change to the achievement of the project goal and impact, and the logical links between them along the results chain. Additionally, the two strands (actor and result) of the theory of change help tackle the attribution issue systematically by incorporating key LFA and OM elements, thus formulating a more comprehensive results chain with explicit logical links between different levels of change.

Case 2 serves as a good example to demonstrate the capability of the synthesis model in accommodating a systematic and tailored approach to addressing the attribution issue through the use of different methods at different levels. In this case, similar mechanisms to those of Case 1 were adopted to address attribution at output and outcome levels through applying LFA elements and tracing the contribution of behavioural and capacity change in intermediary actors to the project goal and impact. However, at impact level, a quasi-experimental design was utilized to solve the attribution issue.

Most of the aforementioned distinctive features of the synthesis model are observations drawn from its practical application in two case study projects originally planned with the LFA. However, the model could also add value to cases which originally adopt OM. As discussed in Chapter 3, in addition to accommodating LFA elements representing relations between outputs and their effects, the model includes goal- and impact-level results (defined in conventional terms and reflecting changes at ultimate beneficiary level), and corresponding indicators to link a project's outcomes with its overall vision in the broader development context.

Through provision of a shared space for key LFA and OM elements, the synthesis model combines the distinctive strengths of both approaches. It integrates the logic of the two and therefore presents a more systematic results chain based on explicit actor orientation. It has the flexibility to stress the LFA element for the purpose of satisfying donor bureaucratic imperatives whilst simultaneously allowing either a comprehensive or tailored evaluation to address specific questions in hand. It represents a critical realist-oriented approach with robust and sophisticated evaluation theory that integrates actor and results focuses and a process of change with various

generative mechanisms and contextual considerations. It thus serves as an alternative option for evaluation in a complex development context.

### **7.3.2 Limitations and challenges arising in the practical application of the synthesis model**

Considerable value addition notwithstanding, there were inevitably some limitations and practical challenges in applying the synthesis model. As my empirical application of the model (as an external and post-implementation evaluation tool) shows, a project theory of change needed to be constructed according to the logic of the model, which then provided a guideline for determining the scope of evaluation – either comprehensive or tailored – and its specific design. Such a process is subject to the critique that I was simply evaluating the project theory I developed rather than the project itself, particularly as I had had no prior involvement in its planning and implementation process.

The evaluation guideline I generated from the constructed theory of change may have been affected by bias and inaccuracy during the reinterpretation and reformulation process on the basis of available data and information (as briefly discussed in Chapter 4). For example, the selection and identification of intermediary actors – boundary partners – in the actor chain is a relatively subjective process and different users of the model may produce a different list of actors. This has implications for future application of the model and is elaborated in following sections. Nevertheless, this does not mean that the model is particularly prone to bias any more than other evaluation approaches. For example, in OM, misspecified boundary partners or outcome challenges are sometimes evaluated, and bias can arise in identifying key elements of the LFA as well.

Additionally, the implementation of the synthesis model requires relevant expertise and experience in both the LFA and OM, and when it is adopted as a holistic framework for comprehensive evaluation this necessitates considerable time, human and financial resources, and technical expertise. The application of the model in both case studies exemplifies these requirements as each required a different evaluation design and technique to review changes at various levels. OM knowledge was essential for evaluating outcome-level results on capacity and behavioural change in OM terms in both cases, and expertise in statistical and experimental design was needed for impact-level evaluation in Case 2. However, a tailored evaluation focusing on some levels of change only would be less demanding.

Another significant challenge is the necessary transformation in mindset from ‘demand’ (recipient/beneficiary) to ‘supply’ (service provider/project implementer). As discussed in Chapter 4, some frontline project staff I interviewed tended to focus mainly on their service delivery and the recipients of their services. They seemed to seldom reflexively think about their own behavioural, capacity or mindset changes, which were also critical in contributing to the achievement of the ultimate goal. This transformation in thinking among various project actors is essential for a better understanding of the fundamental notions and values of the synthesis model, and should be taken into account when applying the model in practice.

Finally, there is the difficulty in aggregating both qualitative and quantitative results gathered with a mixed methods evaluation design aiming to measure different components of the model, particularly when simpler reports of findings are preferred. A few options such as the inductive approach to analysis and quantification of qualitative data were briefly introduced in Chapter 5, but due to some highly context-

dependent and qualitative indicators such as progress markers, the aggregation of data at the final reporting stage remains a challenge. However, this should be less onerous if reporting on one specific level of results in a tailored evaluation with the synthesis model.

## **7.4 Reflections and implications**

In this section, I seek to answer research question 3 through reflecting on the application of the synthesis model, drawing lessons and implications for further research and evaluation practice.

### **7.4.1 Reflection and Lessons from the application of the synthesis model**

First, in this research project, I applied the model as an external and post-implementation evaluation tool and designed the evaluations retrospectively. As discussed in Chapter 4, application of the model at the project end was inevitably affected by many factors such as information and data availability, level of understanding of the project rationale and dynamics on the researcher's part, and the subjective process of reinterpretation and formulation of a new project framework and matrix. However, this research project only exemplifies one way of applying the synthesis model. As with the LFA and OM, to fully utilize the potential of the model, it would be best to adopt it from the project inception and planning stage when many of its actors become involved and take ownership in defining various key elements of the model.

Applying the synthesis model in this manner has certain advantages compared with construction of a theory of change retrospectively during post-implementation evaluation. For example, when progress markers are set up at the project planning

stage through collective effort, some baseline data can be collected and agreed as well. Baseline and subsequent monitoring data reflecting and tracking gradual behavioural changes during the process of project implementation can then be used at the evaluation stage to form the evidence chain. This not only helps reduce bias during evaluation, but also verify and strengthen evaluation results through comparison of baseline and end-line data, which is a more accurate approach than relying on data derived from the recall of past activities retrospectively.

Second, in addition to the importance of timing in terms of synthesis model application, the role and entry point of the evaluator are also critical. Whether the user is a project ‘insider’, a donor-commissioned independent evaluator, or an ‘outsider’ researcher (as in my own position) makes a considerable difference to the standpoint of the evaluator, and the evaluation design and process.

When adopted by an ‘insider’ or project staff member who has deep understanding of project dynamics and access to key project documents, the synthesis model is particularly useful as an internal monitoring and self-evaluation tool designed to achieve both accountability and systematic learning. In these circumstances, evaluation is a more internalized process which is best suited to a tailored approach, and can utilize monitoring data and change stories gathered in continuous reflection sessions during project implementation. There are certain advantages to application of the model by project insider for self-evaluation. For example, it might be easier to gather project stakeholders and obtain a consensus during the process of identifying boundary partners and framing progress markers, which greatly helps ensure monitoring, reflection and evaluation activities. However, the potential issue of lack of independence in such an evaluation role may arise.

A donor-commissioned independent evaluator will find it is relatively less demanding in terms of time and effort in designing evaluation and collecting data if the project has already applied the logic and framework of the synthesis model from the planning stage. An established theory of change can help guide the evaluator in selecting specific data collection method (qualitative, quantitative or mixed methods). However in a project that does not use the synthesis model from the planning stage, the evaluator needs to go through the procedure of reformulating the project framework with the logic of the model. The evaluator can then use this framework to guide the choice to focus either on the whole or part of the theory of change according to the particular evaluation purpose. When conducting an evaluation commissioned by the donor agency, the evaluator has access to key project documents and retains independence while adhering to commissioners' preferences and requirements.

In the non-ideal case where the evaluator-researcher is an outsider and the model is not used systematically from the project's inception, there are difficulties and challenges, as I experienced. In these circumstances, the evaluation design is driven by specific research questions to be answered and the project stage at which the evaluation and research are undertaken, and the data collection process is largely dependent on the availability of relevant project information, existing monitoring and evaluation system, etc. This type of application of the model enables certain evaluator independence, but other issues such as data quality and availability arise.

Third, the donor environment and requirements are also critical factors that influence the application of the synthesis model. Whether the donor agency is open to learning about various outcomes and perspectives and willing to experiment with new

approaches is very important. As revealed in OM section in Chapter 3, a strong donor preference for results-based approaches such as the LFA may limit the recognition of the benefits of OM. The cases of OM in practice I reviewed in Chapter 3 showed that some aid agencies such as Sida have already decided to use it as an alternative to the LFA, or have applied the two approaches to the same project framework. However, this is not common and donor attitude and institutional support remain a key issue in the wider use of OM and its concepts, which in turn potentially impacts on the use of the synthesis model.

For example, in my practical application of the model in two case study projects originally planned with the LFA, different donor emphasis on and interest in evaluation directly influenced the research process. In Case 2, the donor agency was interested in learning about the effects of capacity building. To some extent, this made my empirical research easier as the project team were motivated and tried to be cooperative by allowing me greater access to data than in Case 1.

Finally, the project context under which the model is applied needs to be taken into consideration. Other than the wider economic and political context – as briefly discussed in chapters 5 and 6 in relation to each case study – institutional setting, status and relationship at organizational level are also key factors, particularly given the actor-centred focus of the synthesis model. For example, in Case 1, a key actor in both the lead project team and boundary partner groups – local consumer associations – transpired to be government agencies rather than civil society organizations. Advance research to gather such knowledge would help to better understand the role of actors and their relationships with each other when drawing boundaries between different groups in the actor chain in order to formulate a feasible and effective actor



strand of the theory of change.

#### **7.4.2 Appropriate circumstances for using the synthesis model**

In this research project, I applied the synthesis model in two aid projects which covered different sectors and areas such as sustainable consumption, poverty reduction and micro finance. As discussed in earlier chapters, they both had explicit behavioural change and capacity building elements. However, this does not mean that the synthesis model can only be used in projects which include these elements – it is applicable to those that could feasibly adopt either the LFA or OM. This notion accords with the conviction of the original authors of the model that the LFA and OM are not mutually exclusive but complementary, and that a synthesized model could integrate the strengths of both approaches. Both case study projects originally used the LFA for planning, monitoring and self-evaluation, which helped make the application of the synthesis model feasible. Given the dominant status of the LFA, there should be considerable potential for using such a synthesis model.

As mentioned in introducing the synthesis model, its form remains open and leaves space for users to emphasize either the LFA or OM elements according to the requirements of a specific project. It was also noted that the model can be used in a tailored or a comprehensive manner with varied focus and scope, in a simple linear fashion utilizing LFA logic, or as a more sophisticated holistic package. For example, if there is good knowledge of contextual factors and project goals can be easily agreed and measured, the synthesis model can probably be applied in a tailored and simple form emphasizing LFA elements as in conventional evaluation. In a more complex project involving multiple groups of actors and a multi-level institutional structure and cooperation mechanism, the model may need to emphasize its OM elements to

help differentiate actors and their relationships in order to construct a more inclusive and feasible theory of change.

As reflected by the evaluation cases utilizing the LFA and OM in Chapter 3, various types of projects apply these two approaches, including education, agriculture, budget support, networking and advocacy initiatives, at either country or community level. The synthesis model is considered to be suitable for these projects as well as a wide range of interventions which seek to bring about change, recognize the importance of people, and emphasize quality, process and the notion of contribution, in addition to other valued aspects such as focus on results. Thus, for example, the model can be used when there is a need to establish an explicit link between process and outcome evaluation.

As argued earlier, each evaluation approach has its niche and is framed to serve a different purpose and scope. A rigorous approach should be well situated to the evaluation subject, purpose and questions in a particular context with its various real-world constraints. In addition to the 'mix and match' type of combined use of different tools such as the LFA and OM (as mentioned in Chapter 3), and provided evaluators appreciate the range of approaches currently available and make a careful informed choice in each given case, the synthesis model can enrich the evaluation toolbox as an additional option. However, to realize the full potential of the model, careful study and practical consideration of the project circumstances is necessary to determine whether it appears to be a feasible option.

#### **7.4.3 Further areas for research on the synthesis model**

The research design, process and findings presented in this thesis are confined to the

limited scope of a PhD project, and the knowledge and experiences of the researcher. There are some potential research areas I was not able to explore, such as a deeper treatment of synthesis model assumptions than that provided in chapters 5 and 6, and the value of other elements, including organizational practice and strategy maps in terms of achieving institutional learning and developing tailored support strategies for concerned actors. Inclusion of these components would entail in-depth involvement in the project from the planning stage and also continuous reflection during project implementation.

Finally, insights afforded by this research project with two different case-specific examples of application of the synthesis model provide references and help future ‘under-similar-conditions’ usage of the model. Based on this, it would be helpful to accumulate clusters of cases and extend the application of the model to diverse types of projects to generate further insight into its applicability and best usage. Moreover, given the critical influence of donor recognition of a particular approach, actor-centred research on donor receptivity and perceptions of the model is also worth pursuing in future applications to gather donor experiences and insights for further possible adaptations of the model.

## Bibliography

Adams, J. (2007) *Learning and Accountability: A Monitoring & Evaluation Consultant's Perspective*, INTRAC Praxis Note 32, Oxford, UK: International NGO Training and Research Centre (INTRAC).

Ambrose, K. and Roduner, D. (2009) *A conceptual fusion of the logical framework approach and outcome mapping*, OM ideas Paper 1, May 2009. [Online]. Available at: <http://www.outcomemapping.ca/resource/resource.php?id=231> [Accessed 23 July 2017].

Argyris, C. and Schön, D.A. (1978) *Organizational Learning: A Theory of Action Perspective*. Massachusetts: Addison-Wesley.

Armstrong, J., Carden, F., Coe, A. and Earl, S. (2000) *IMFNS: Outcomes Assessment*. Ottawa: Governance Network and the International Development Research Centre (IDRC).

Armytage, L. (2011) Evaluating aid: An adolescent domain of practice. *Evaluation*, 17(3): 261-276.

Asian Development Bank (ADB) (2013) *Cost-benefit analysis for development: A Practical Guide*. Mandaluyong City: Asian Development Bank.

Astbury, B. and Leeuw, F.L. (2010) Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation. *American Journal of Evaluation*, 31(3): 363-381.

AusAID (2005) *AusGuideline: The Logical Framework Approach*. Canberra: Australian Agency for International Development.

AusAID (2013) *Independent Review of the Pacific Horticultural and Agricultural Market Access (PHAMA) Program*. Final report, AusAID Pacific Division.

Bakewell, O. and Garbutt, A. (2005) *The use and abuse of the logical framework approach: A review of International Development NGOs Experiences*. Oxford, UK: INTRAC.

Bamberger, M., Rao, V. and Woolcock, M. (2009) *Using mixed methods in monitoring and evaluation: experiences from international development*. BWPI Working Paper 107, Manchester: Brooks World Poverty Institute (BWPI), University of Manchester.

Bamberger, M. (2012) *Introduction to mixed methods in impact evaluation*. Impact Evaluation Notes 3. Washington DC: InterAction and Rockefeller Foundation.

Bamberger, M., Rugh, J. and Mabry, L. (2012) *RealWorld Evaluation: Working Under Budget, Time, Data, and Political Constraints*, Second Edition. London: Sage.

Barnett, C., Barr, J., Christie, A., Duff, B. and Hext, S. (2010) *Measuring the Impact and Value for Money of Governance & Conflict Programmes*. Final Report. [Online]. Available at: [http://www.dfid.gov.uk/r4d/pdf/outputs/mis\\_spc/60797\\_itad-vfm-report-](http://www.dfid.gov.uk/r4d/pdf/outputs/mis_spc/60797_itad-vfm-report-)

[dec10.pdf](#) [Accessed 16 May 2017].

Barnett, C. and Gregorowski, R. (2013) *Learning about Theories of Change for the Monitoring and Evaluation of Research Uptake*. IDS Practice Paper Brief 14, Brighton: Institute of Development Studies (IDS).

BetterEvaluation (2014) *Cost-Benefit Analysis*. [Online]. Available at: <http://www.betterevaluation.org/en/evaluation-options/CostBenefitAnalysis> [Accessed 27 July 2017].

Bloom, H. (2006) *The Core Analytics of Randomized Experiments for Social Research*. MDRC Working Papers on Research Methodology, New York: Manpower Demonstration Research Corporation (MDRC).

Bornstein, L. (2003) Management Standards and Development Practice in the South Africa Aid Chain. *Public Administration and Development*, 23(5): 393-404.

Bortcosh, S. and Gibby, P. (2016) *Assessing and Managing Value for Money: Lessons for NGOs*. Summary Paper, Mango and the Value-for-Money Learning Group of the Programme Partnership Arrangements (PPA) Learning Partnership.

BUCEA (2012) *Sustainable Consumption\_Urban\_China 2012 Interim Narrative Report*. Beijing: Beijing University of Civil Engineering and Architecture (BUCEA).

BUCEA (2013) *Sustainable Consumption\_Urban\_China 2013 Interim Narrative Report*. Beijing: Beijing University of Civil Engineering and Architecture (BUCEA).

BUCEA (2015) *Sustainable Consumption\_Urban\_China Final Narrative Report*. Beijing: Beijing University of Civil Engineering and Architecture (BUCEA).

CARE (2017) *Pathways to Empowerment*. [Online]. Available at: <http://www.carepathwaystoempowerment.org/resources-2/> [Accessed 27 July 2017].

Cartwright, N. (2007) Are RCTs the Gold Standard? *BioSocieties*, 2(1): 11-20.

Cartwright, N. and Munro, E. (2010) The limitations of randomized controlled trials in predicting effectiveness. *Journal of Evaluation in Clinical Practice*, 16(2): 260-266.

CEDIL (2018) *Innovative Methods for Useful Results*. [Online]. Available at: <https://cedilprogramme.org> [Accessed 20 June 2018].

Centre for Theory of Change (2016) *How does theory of change work?* [Online]. Available at: <http://www.theoryofchange.org/what-is-theory-of-change/how-does-theory-of-change-work/#3> [Accessed 27 July 2017].

Chambers, R., Karlan, D., Ravallion, M. and Rogers, P. (2009) *Designing impact evaluations: different perspectives*. 3ie Working Paper 4, New Delhi: International Initiative for Impact Evaluation (3ie).

Chen, H.T. and Rossi, P.H. (1987) The theory-driven approach to validity. *Evaluation*

and *Program Planning*, 10(1): 95-103.

Cilliers, P. (2001) Boundaries, Hierarchies and Networks in Complex Systems. *International Journal of Innovation Management*, 5(2): 135–147.

Clarke, P. and Taylor, P. (2008) *Capacity for a change*. Document based on outcomes of the ‘Capacity Collective’ workshop Dunford House, 25-27 September 2007, Brighton: Institute of Development Studies (IDS).

Conlin, S. and Stirrat, R.L. (2008) Current Challenges in Development Evaluation. *Evaluation*, 14(2): 193 – 208.

Colombo, M. (2017) Experimental Philosophy of Explanation Rising: The Case for a Plurality of Concepts of Explanation. *Cognitive Science*, 41(2): 503-517.

Colombo, M., Bucher, L. and Sprenger, J. (2017) Determinants of Judgments of Explanatory Power: Credibility, Generality, and Statistical Relevance. *Frontiers in Psychology*, 8:1430.

Coryn, C.L.S., Noakes, L.A., Westine, C.D., Schröter, D.C. (2011) A systematic review of theory-driven evaluation practice from 1990 to 2009. *American Journal of Evaluation*, 32(2): 199-226.

Couillard, J., Garon, S., and Riznic, G. (2009) The logical framework approach millennium. *Project Management Journal*, 40(4): 31-44.

Cracknell, B.E. (2000) *Evaluating Development Aid: Issues, Problems and Solutions*. London: Sage.

Crawford, P., Perryman, J. and Petocz, P. (2004) Synthetic indices: a method for evaluating aid project effectiveness. *Evaluation*, 10(2): 175-192.

Crawford, P., Swete, D., Mitchell, C. and Bryce, P. (2005) *Aristotle and Plato at it again? Philosophical divergence within international aid project design and monitoring & evaluation*. [Online]. Available at: [https://www.outcomemapping.ca/download/kaia\\_ambroseyahoo.ca\\_en\\_070105\\_Aristotle%20and%20Plato%20at%20it%20again.pdf](https://www.outcomemapping.ca/download/kaia_ambroseyahoo.ca_en_070105_Aristotle%20and%20Plato%20at%20it%20again.pdf) [Accessed 23 July 2017].

Dale, R. (2003) The logical framework: an easy escape, a straightjacket, or a useful planning tool? *Development in Practice*, 13(1): 57-70.

Dalkin, S.M., Greenhalgh, J., Jones, D., Cunningham, B., and Lhussier, M. (2015) What’s in a mechanism? Development of a key concept in realist evaluation. *Implementation Science*, 10(1): 1-7.

Danermark, B., Ekstrom, M., Jakobsen, L., and Karlsson, J.C. (2002) *Explaining Society: Critical Realism in the Social Sciences*. Milton Park, New York: Psychology Press.

Davies, R. (2004) Scale, complexity and the representation of the theories of change.

*Evaluation*, 10(1): 101-121.

Davies, R. (2005) Scale, complexity and the representation of theories of change: Part II. *Evaluation*, 11(2): 133-149.

Delahais, T. and Toulemonde, J. (2012) Applying contribution analysis: Lessons from five years of practice. *Evaluation*, 18(3): 281–293.

DFID (2006) *Promoting financial services for poverty reduction in Bangladesh (PROSPER)*. Project document, Department for International Development (DFID).

DFID (2009) *Political Economy Analysis How to Note*. A Practice Paper, London: Department for International Development (DFID).

DFID (2011) *How to note: Guidance on using the revised logical framework*. A Practice Paper, London: Department for International Development (DFID).

DFID (2013) *International Development Evaluation Policy*. London: Department for International Development (DFID).

DFID (2014) *DFID Evaluation Strategy 2014-2019*. London: Department for International Development (DFID).

Duflo, E., Glennerster, R., and Kremer, M. (2006) *Using randomization in development economics research: A toolkit*. Technical Working Paper 333, Cambridge, US: National Bureau of Economic Research.

Dyer, K. (2011) *Think piece: Working with log-frames and outcome mapping in the context of the accountability Tanzania programme (AcT)*. [Online]. Available at: <http://www.accountability.or.tz/wp-content/uploads/2011/03/Think-Piece-on-OM.pdf> [Accessed 3 February 2014].

Earl, S., Carden, F. and Smutylo, T. (2000) *Outcome mapping: focusing monitoring and evaluation of development programs on changes in partners*. Selected Proceedings from a Seminar and Workshop organized by the Development Bank of Southern Africa, the African Development Bank and the World Bank, Johannesburg, 25–29 September 2000.

Earl, S., Carden, F. and Smutylo, T. (2001) *Outcome Mapping: Building Learning and Reflection into Development Programs*. Ottawa: IDRC.

Earle, L. (2003) *Lost in the Matrix: The Logframe and the Local Picture*. Paper for the 5th INTRAC Evaluation Conference: Measurement, Management and Accountability? the Netherlands, 31 March – 4 April 2003.

Ebrahim, A. (2005) Accountability Myopia: Losing Sight of Organizational Learning. *Nonprofit and Voluntary Sector Quarterly*, 34 (1): 56-87.

ECODE (2011) *Critical Study of the Logical Framework Approach in the Basque Country*. Bilbao: Estudio de Cooperacion al Desarrollo (ECODE) – Development Cooperation Study.

Enfield, S. and Forsberg, L. (2012) *Interim Evaluation of Support to the Church of Sweden's Community Based Psychosocial Support in Humanitarian Assistance 2006–2011*. Sida Decentralised Evaluation 2012:1, Sweden: Swedish International Development Cooperation Agency (Sida).

Ergas, H. (2009) In Defence of Cost-Benefit Analysis. *Agenda: A Journal of Policy Analysis and Reform*, 16 (3): 31-40.

EuropeAid (2012) *ROM Handbook: Results-oriented Monitoring*. EuropeAid EVA/129-719. [Online]. Available at: <http://capacity4dev.ec.europa.eu/rom> [Accessed 6 January 2017].

FASID (2005) *New Approaches to Development and Changing Sector Issues: Recent Developments in Aid Evaluation*. Tokyo: Foundation for Advanced Studies on International Development (FASID).

Fournier, D.M. (2005) 'Evaluation', [in] Mathison, S. (ed.) *Encyclopedia of Evaluation*. Thousand Oaks, CA: Sage, pp.139-140.

Fritz, V., Levy, B. and Ort, R. (2014) *Problem-driven Political Economy Analysis: The World Bank's Experience*. Washington DC: World Bank.

Funnell, S. C. and Rogers, P.J. (2011) *Purposeful program theory: effective use of theories of change and logic models*. San Francisco: Jossey-Bass/Wiley.

Gasper, D. (1997) *Logical frameworks: A critical assessment managerial theory, pluralistic practices*. Working paper 264, the Hague: Institute of Social Studies (ISS).

Gasper, D. (2000) Evaluating the logical framework approach 'towards learning-oriented development evaluation'. *Public Administration and Development*, 20(1): 17-28.

Gates, E. and Dyson, L. (2017) Implications of the Changing Conversation About Causality for Evaluators. *American Journal of Evaluation*, 38(1): 29-46.

Gerring, J. (2007) *Case Study Research: Principles and Practices*. Cambridge: Cambridge University Press.

Grove, N.J. and Zwi, A.B. (2008) Beyond the log frame: a new tool for examining health and peacebuilding initiatives. *Development in Practice*, 18(1): 66-81.

Guba, E.G. and Lincoln, Y.S. (1989) *Fourth Generation Evaluation*. Newbury Park, CA: Sage.

Guijt, I. (2008) *Seeking Surprise: Rethinking monitoring for collective learning in rural resource management*. PhD thesis. Wageningen: Wageningen University

Guijt, I. (2010) 'Exploding the myth of incompatibility between accountability and learning', [in] Ubels, J., Acquaye-Baddoo, N. and Fowler, A. (ed.) *Capacity*



*development in practice*. London: Earthscan, pp. 277-291.

Hansen, H.F. (2005) Choosing Evaluation Models: A Discussion on Evaluation Design. *Evaluation*, 11(4): 447–462.

Harley, K. (2005) Learning from logframes: reflections on three educational development projects in East and Southern Africa. *Compare: A Journal of Comparative and International Education*, 35 (1): 27-42.

Hart, C. (2005) *Doing Your Masters Dissertation*. Sage Essential Study Skills, London: Sage.

Hearn, S. and Ambrose, K. (2013) *Introduction to Outcome Mapping*. Presentation at the American Evaluation Association Conference, Washington DC, October 2013.

Hockley, N. (2014) Cost–benefit analysis: a decision-support tool or a venue for contesting ecosystem knowledge? *Environment and Planning C: Politics and Space*, 32 (2): 283-300.

Howard, G., Jeger, M. and Wilson-Grau, R. (2011) *BioNET 2007-2010 Evaluation: An outcomes evaluation and assessment of the prospects for BioNET to increase its impact on food security, in particular through greater support to plant health systems*. [Online]. Available at: <https://www.outcomemapping.ca/download/BioNET-Global-Programme-Evaluation.pdf> [Accessed 26 July 2017].

Hudson, D. and Leftwich, A. (2014) *From political economy to political analysis*. DLP Research Paper 25, Birmingham: Developmental Leadership Program (DLP), International Development Department, University of Birmingham.

Hummelbrunner, R. (2010) ‘Beyond logframe: critique, variations and alternatives’, [in] Fujita, N. (ed.) *Beyond Logframe: Using Systems Concepts in Evaluation*. Tokyo: Foundation for Advanced Studies on International Development (FASID), pp.1-33.

ICAI (2011) *ICAI’s Approach to effectiveness and value for money*. Report 1, London: Independent Commission for Aid Impact (ICAI).

IDRC (2010) *Outcome Mapping and the Logical Framework Approach: Can they share a space?* Ottawa: International Development Research Centre (IDRC). [Online]. Available at: <https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/50778/IDL-50778.pdf?sequence=1&isAllowed=y> [Accessed 25 July 2017].

INTRAC (2012) *Theory of Change: what’s it all about?* INTRAC newsletter 51, Oxford, UK: International NGO Training and Research Centre (INTRAC).

Jackson, P. (2012) *Value for money and international development: Deconstructing myths to promote a more constructive discussion*. Paris: OECD.

Jonas, N., Jonas, H., Steer, L. and Datta, A. (2009) *Improving Impact Evaluation production and use*. Working paper 300, London: Overseas Development Institute (ODI).

Jones, H. (2006) *Making Outcome Mapping work: evolving experiences from around the world*. Ottawa: Outcome Mapping Learning Community.

Jones, H. and Hearn, S. (2009) *Outcome Mapping: a realistic alternative for planning, monitoring and evaluation*. Background note, London: Overseas Development Institute (ODI).

Keystone (2009) *Developing a theory of change: A guide to developing a theory of change as a framework for inclusive dialogue, learning and accountability for social impact*. IPAL GUIDE 2, Impact Planning, Assessment and Learning. [Online]. Available at: <https://keystoneaccountability.org/wp-content/uploads/files/2%20Developing%20a%20theory%20of%20change.pdf> [Accessed 25 July 2017].

Khalily, M.A., Latif, M.A., Hasan, M.M., Nargis, F., Akhter, N., Muneer, F., Khaleque, A., Begum, A., Islam, S., Rabbani, A., and Roy, S. (2013) *Impact of PRIME interventions at the household level – 5th round*. Dhaka: Institute of Microfinance.

Khalily, M.A., Hasan, M., Akhter, N. (2016) *A longitudinal study to assess the impact of PRIME in Greater Rangpur: 7th Round*. Dhaka: Institute of Microfinance.

Kilby, P. (2004) Is empowerment possible under New Public Management environment? Some lessons from India. *International Public Management Journal*, 7(2): 207-225.

Kishekya, D., Mauremootoo, J. and Smith, R.D. (2013) *An outcomes evaluation of Search For Common Ground's The Team Tanzania*. [Online]. Available at: <http://www.outcomemapping.ca/resource/an-outcomes-evaluation-of-search-for-common-grounds-the-team-tanzania> [Accessed 25 July 2017].

Kogen, L. (2017) What have we learned here? Questioning accountability in aid policy and practice. *Evaluation*, 24(1): 98–112.

Kolb, D.A. (1984) *Experiential Learning: Experience as the Source of Learning and Development*. Englewood Cliffs, NJ: Prentice-Hall.

Kontinen, T. (2010) 'Evaluation in NGO development cooperation: Room for communities of inquiry?', [in] Moksnes, H. and Melin, M. (ed.) *Power to the People? (Con-)Tested Civil Society in Search of Democracy*. Uppsala: Uppsala Centre for Sustainable Development, pp.230-234.

Kothari, U. (2000) *Developing guidelines for assessing achievement in the eight focal areas of social development and for assessing outcomes*. Social Development Systems for Coordinated Poverty Eradication (SD SCOPE) Paper 10, Manchester: Institute for Development Policy and Management, University of Manchester.

Lampis, A. (2005) 'Exploring the temporal logic model: A Colombian case study evaluation internally displaced people assistance', [in] Campbell, J. and Holland, J. (Eds.) *Methods in development research: combining qualitative and quantitative*

*approaches*. London: ITDG Publications, pp.177-190.

Landoni, P. and Corti, B. (2011) The management of international development projects: Moving toward a standard approach or differentiation? *Project Management Journal*, 42(3): 45-61.

Lanser, P., Dom, C., Orivel, F. and Ouédraogo, J. P. (2006) *Joint Evaluation of General Budget Support 1994-2004: Burkina Faso Country Report*. Glasgow: Evaluation Department, Department for International Development (DFID).

Lehtonen, M. (2005) OECD Environmental Performance Review Programme: Accountability (f)or Learning? *Evaluation*, 11(2): 169-188.

Lennie, J. and Tacchi, J. (2014) Bridging the divide between upward accountability and learning-based approaches to development evaluation: Strategies for an enabling environment. *Evaluation Journal of Australasia*, 14(1): 12-23

Leong, W.Y. and Lim Y.L. (2016) 'Cost-Benefit Analysis', [in] Shroff et al. (ed.) *Evidence-based Policymaking in Singapore: A policymaker's toolkit*. Singapore: Civil Service College, pp.47-55.

Levinthal, D.A. and March, J.G. (1993) The myopia of learning. *Strategic Management Journal*, 14: 95-112.

Long, N. (2001) *Development sociology: actor perspectives*. London and New York: Routledge.

Mayne, J. (2001) Addressing attribution through contribution analysis: using performance measures sensibly. *Canadian Journal of Program Evaluation*, 16 (1): 1-24.

Mayne, J. (2008) *Contribution Analysis: An Approach to Exploring Cause and Effect*. ILAC Brief 16, Rome: Institutional Learning and Change Initiative (ILAC).

Mayne, J. (2012a) Contribution analysis: Coming of age? *Evaluation*, 18(3): 270-280.

Mayne, J. (2012b) *Making causal claims*. ILAC Brief 26, Rome: Institutional Learning and Change Initiative (ILAC).

McLoughlin, C. (2014) *Political economy analysis: Topic guide*, 2nd edition. Birmingham: GSDRC, University of Birmingham.

Meadows, D., Randers, J., and Meadows, D. (2005) Tools for the transition to sustainability. *The Systems Thinker*, 16(9): 2-9.

Mikkelsen, B. (2005) *Methods for development work and research: a new guide for practitioners*. New Delhi: Sage

Minsky, M. (1988) *The Society of Mind*. New York: Simon and Schuster.

Morra Imas, L.G. and Rist, R.C. (2009) *The road to results: Designing and conducting*

*effective development evaluations*. Washington, DC: World Bank.

Moses, J.W. and Knutsen, T.L. (2007) *Ways of Knowing: Competing Methodologies in Social and Political Research*. New York: Palgrave macmillan.

ODI (2018) *The Methods Lab: improving practice for impact evaluation in DFAT*. [Online]. Available at: <https://www.odi.org/projects/2599-methods-lab-improving-practice-impact-evaluation-dfat> [Accessed 20 June 2018].

OECD-DAC (1991) *Principles for Evaluation of Development Assistance*. Paris: OECD.

OECD (2000) *Result-based management in the development cooperation agencies: a review of experiences*. Background report, Paris: OECD.

OECD-DAC (2001) *Evaluation Feedback for Effective Learning and Accountability*. Paris: OECD.

OECD-DAC (2002) *Glossary of Key Terms in Evaluation and Results Based Management*. Paris: OECD.

OECD-DAC (2006) *Cost-benefit analysis and the environment: recent developments*. Paris: OECD.

OECD (2014) *Measuring and managing results in development co-operation: A review of challenges and practices among DAC members and observers*. Paris: OECD.

OECD (2016) *Evaluation Systems in Development Co-operation: 2016 Review*. Paris: OECD.

Patton, M.Q. (2002) *Qualitative Research and Evaluation Methods*, 3rd edition. Thousand Oaks, CA: SAGE.

Patton, M.Q. (2008) *Utilization-Focused Evaluation* 4th edition. Thousand Oaks, CA: Sage.

Patton, M.Q. (2011) 'Utilization-Focused Evaluation: Real-Time and Prospective Aspects' in *Exploring the Potential of Real-Time and Prospective Evaluations*. Evaluation brief, Washington DC: World Bank.

Patton, M.Q. (2012) A utilization-focused approach to contribution analysis. *Evaluation*, 18(3): 364-377.

Pawson, R. and Tilley, N. (1997) *Realistic Evaluation*, London: Sage.

Pettit, J. and Acosta, A.M. (2014) Power above and below the waterline: bridging political economy and power analysis. *IDS Bulletin*, 45(5): 9-22.

Picciotto, R. (2002) *The logic of renewal: Evaluation at the World Bank (1992-2002)*. OED Working Paper 82088, Washington DC: Operations Evaluation Department

(OED), World Bank.

Picciotto, R. (2007) The New Environment for Development Evaluation. *American Journal of Evaluation*, 28(4): 509-521.

PKSF (2013) *PRIME Result-Based Monitoring (RBM): Results for 2011 and 2012*. Dhaka: Palli Karma Sahayak Foundation.

PKSF (2017) *Programmed Initiatives for Monga Eradication (PRIME)*. [Online]. Available at: [http://pkssf-bd.org/portal/web/?page\\_id=122](http://pkssf-bd.org/portal/web/?page_id=122) [Accessed 25 July 2017].

Poole, A. (2011) *Political economy assessments at sector and project levels*. How-To Note, Washington DC: World Bank.

Powell, S., Celebivic, I., Bratovic, E. and Sisic, A. (2008) *Outcome Mapping Evaluation of Six Civil Society Projects in Bosnia and Herzegovina*. Sida Evaluation 2008:17, Summary Report, Sweden: Swedish International Development Cooperation Agency (Sida).

Preskill, H. (2008) Evaluation's Second Act: A Spotlight on Learning. *American Journal of Evaluation*, 29(2): 127-138

Ravallion, M. (2009) Evaluation in the Practice of Development. *The World Bank Research Observer*, 24(1): 29-53.

Regeer, B.J., de Wildt-Liesveld, R., van Mierlo, B. and Bunders, J.F.G. (2016) Exploring ways to reconcile accountability and learning in the evaluation of niche experiments. *Evaluation*, 22(1): 6-28.

Reich, M.R., and Balarajan, Y. (2012) *Political Economy Analysis for Food and Nutrition Security*. Health, Nutrition and Population (HNP) discussion paper, Washington DC: World Bank.

Reinertsen, H., Bjørkdahl, K. and McNeill, D. (2017) Confronting the Contradiction - An exploration into the dual purpose of accountability and learning in aid evaluation. Expert Group on Aid Studies (EBA). [Online]. Available at: <http://eba.se/wp-content/uploads/2017/05/Webbversion-FINAL-blanksida.pdf>. [Accessed 17 April 2018]

Roduner, D., Schlappi, W. and Egli, W. (2008) Logical Framework Approach and Outcome Mapping: A constructive attempt of synthesis. *Rural Development News* 2/2008.

Roebeling, G., Oomes, I., Alpenidze, I. and Zuijderduijn, M. (2014) *Independent Mid-Term Review of the Strategic Framework 2011-2015 of the International Land Coalition and Management Response*. International Land Coalition.

Rossi, P.H. (2004) 'My views of evaluation and their origins', [in] Alkin, M.C. (ed.) *Evaluation roots: Tracing theorists' views and influences*. Thousand Oaks, CA: Sage, pp.122-131.

Rossi, P.H., Lipsey, M.W., and Freeman, H.E. (2004) *Evaluation: A systematic approach*, 7th edition, Thousand Oaks, CA: Sage.

Sayer, A. (2000) *Realism and Social Science*. London: Sage.

Schillemans, T. (2008) Accountability in the Shadow of Hierarchy: The Horizontal Accountability of Agencies, *Public Organization Review*, 8(2): 175-194.

Scriven, M. (1996) The theory behind practical evaluation. *Evaluation*, 2(4): 393-404.

Scriven, M. (2008) A Summative Evaluation of RCT Methodology: & An Alternative Approach to Causal Research. *Journal of MultiDisciplinary Evaluation*, 5 (9): 11-24.

SDC (2011) *Planning and monitoring in results-based management of projects and programmes*. Training course script, Bern: Swiss Agency for Development and Cooperation (SDC)

Segone, M. (2016) 'A call for action to leave no one behind', *European Evaluation Society (EES) Newsletter*, February 2016, p.3.

Shadish, W.R., Cook, T.D. and Leviton, L.C. (1991) *Foundations of Program Evaluation: Theories of Practice*. Newbury Park, CA: Sage.

Shadish, W.R. and Luellen, J.K. (2005) 'History of Evaluation', [in] Mathison, S. (ed.) *Encyclopedia of Evaluation*. Thousand Oaks, CA: Sage, pp.184-187.

Shadish, W.R., Clark, M.H., and Steiner, P.M. (2008) Can nonrandomized experiments yield accurate answers? A randomized experiment comparing random and nonrandom assignments. *Journal of the American Statistical Association*, 103(484): 1334-1344

Shapiro, S. and Morrall III, J. F. (2012) The triumph of regulatory politics: benefit–cost analysis and political salience. *Regulation & Governance*, 6 (2): 189-206.

Shaxson, L. and Clench, B. (2011) *Outcome mapping and social frameworks-tools for designing, delivering and monitoring policy via distributed partnerships*, Delta Partnership Working Paper 1. [Online]. Available at: [https://www.outcomemapping.ca/download/simonhearn\\_en\\_Working%20paper%20-%20outcome%20mapping%20%20social%20frameworks.pdf](https://www.outcomemapping.ca/download/simonhearn_en_Working%20paper%20-%20outcome%20mapping%20%20social%20frameworks.pdf) [Accessed 25 July 2017].

Sheriff, N. and Schuetz, T. (2010) *Benefits and Challenges of Applying Outcome Mapping in an R4D project*. CBFC (Community-based Fish Culture) Working Paper 6 (draft), Penang, Malaysia: WorldFish Center.

Sida (2004) *The Logical Framework Approach: A summary of the theory behind the LFA method*. Sweden: Swedish International Development Cooperation Agency (Sida).

Smith, R., Mauremootoo, J. and Rassmann, K. (2012) *Ten years of Outcome Mapping adaptations and support*. Outcome Mapping Learning Community.

Smutylo, T. (2001) *Crouching impact, hidden attribution: overcoming threats to*

*learning in development programs*. Draft Learning Methodology Paper prepared for the Block Island Workshop on Across Portfolio Learning, 22–24 May 2001. Ottawa: International Development Research Centre (IDRC). [Online]. Available at: [https://www.outcomemapping.ca/download/admin\\_en\\_Crouching\\_Impact\\_Hidden\\_Attribution.pdf](https://www.outcomemapping.ca/download/admin_en_Crouching_Impact_Hidden_Attribution.pdf) [Accessed 25 July 2017].

Solmeyer, A.R. and Constance, N. (2015) Unpacking the “Black Box” of Social Programs and Policies: Introduction. *American Journal of Evaluation*, 36(4): 470-474

Stein, D. and Valters, C. (2012) *Understanding ‘Theory Of Change’ in International Development: A Review Of Existing Knowledge*. London: Justice and Security Research Programme, London School of Economics and Political Science (LSE).

Stern, E., Stame, N., Mayne J., Forss, K., Davies, R. and Befani, B. (2012) *Broadening the range of designs and methods for impact evaluations*. DFID Working Paper 38, London: Department for International Development (DFID).

Sumner, A. and Tribe, M.A. (2008) *International Development Studies: Theories and Methods in Research and Practice*. London: Sage.

SWITCH-Asia (2015a) *Implementing Sustainable Consumption in Civil Society of Urban China*. Project Information Sheet, SWITCH-Asia Programme launched by the European Commission.

SWITCH-Asia (2015b) *Implementing Sustainable Consumption in Civil Society of Urban China*. Project Impact Sheet, SWITCH-Asia Programme launched by the European Commission.

Teddle, C. and Tashakkori, A. (2009) *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioural Sciences*. Thousand Oaks, CA: Sage.

Thomas, D.R. (2006) General Inductive Approach for Analyzing Qualitative Evaluation Data. *American Journal of Evaluation*, 27(2): 237-246.

Thomas, V. (2010) *Evaluation for Greater Development Effectiveness*. Paper prepared for the Korea Development Institute Conference on Economic Development and Impact Evaluation, Seoul, South Korea, 15-16 November 2010. [Online]. Available at: [http://siteresources.worldbank.org/EXTDIRGEN/Resources/evaluation\\_thomas\\_KDI.pdf](http://siteresources.worldbank.org/EXTDIRGEN/Resources/evaluation_thomas_KDI.pdf) [Accessed 26 July 2017].

Toulemonde, J. (2016) ‘Is there a trade-off between accountability and learning in evaluation?’ *European Evaluation Society (EES) Newsletter*, February 2016, p.12.

UNDP (2008) *Capacity Development*. Practice Note, New York: UNDP.

UNDP (2009) *Handbook on planning, monitoring and evaluating for development results*. New York: UNDP.

United Nations (2015) *Transforming our World: The 2030 Agenda for Sustainable*

*Development*. New York: United Nations.

USAID (2016) *USAID Evaluation Policy: Evaluation Learning from Experience*. Washington DC: U.S. Agency for International Development (USAID).

Valters, C., Cummings, C. and Nixon, H. (2016) *Putting learning at the centre: Adaptive development programming in practice*. ODI Report, London: Overseas Development Institute (ODI).

Van Der Meer, F.B. and Edelenbos, J. (2006) Evaluation in Multi-Actor Policy Processes: Accountability, Learning and Co-operation. *Evaluation*, 12(2): 201-218.

Van Ongevalle, J., Chipimbi, R. and Sibanda, M. (2009) *Analysing Outcome Mapping Monitoring Data: The case of the Quality Education and Vulnerability Programme*. Working Paper. [Online]. Available at: [https://www.outcomemapping.ca/download/janvozol.co.zw\\_en\\_Analysis%20of%20OM%20data%20-%20VVOB%20Zimbabwe.pdf](https://www.outcomemapping.ca/download/janvozol.co.zw_en_Analysis%20of%20OM%20data%20-%20VVOB%20Zimbabwe.pdf) [Accessed 27 July 2017].

Van Ongevalle, J., Huyse, H., Temmink, C., Boutylkova, E., Maarse, A. (2012) *Dealing with complexity through “actor-focused” Planning, Monitoring & Evaluation (PME): From results-based management towards results-based learning*. Final output of the PSO Thematic Learning Programme (TLP) on Planning, Monitoring and Evaluation (PME) of Complex Processes of Social Change, The Hague: PSO (An association of sixty Dutch development organisations).

Van Ongevalle, J. (2013) *Outcome Mapping: from alternative to mainstream? The OM usefulness barometer*. Keynote address at OM Master Class, Brussels, Belgium, 25-26 March 2013. [Online]. Available from: <http://www.outcomemapping.ca/resource/resource.php?id=413> [Accessed 27 July 2017]

Van Ongevalle, J., Huyse, H., and Van Petegem, P. (2014) Dealing with complexity through actor-focused planning, monitoring and evaluation (PME). *Evaluation*, 20(4): 447-466.

Van Ongevalle, J. and Fonteneau, B. (2014) *Learning about the effects of development education programmes: Towards a learning centred monitoring and evaluation practice*. Leuven: HIVA – Research Institute for Work and Society.

Vedung, E. (2010) Four waves of evaluation diffusion. *Evaluation*, 16(3): 263-277.

Vogel, I. (2012) *Review of the Use of ‘Theory of Change’ in International Development*. Review Report, London: Department for International Development (DFID).

Webb, L.D. (1976) Cost-Benefit Analysis: An Accountability Asset. *Journal of Education Finance*, 2(2): 209-223.

Weiss, C.H. (1995) ‘Nothing as Practical as a Good Theory: Exploring Theory-based Evaluation in Complex Community Initiatives for Children and Families’, [in] Connell, J., Kubish, A., Schorr, L. and Weiss, C. (ed.) *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts*. Washington DC: Aspen Institute, pp.65-



92.

Weiss, C.H. (1997) Theory-based evaluation: Past, present, and future. *New directions for evaluation*, 1997(76): 41-55.

Weiss, C.H. (2000) Which links in which theories shall we evaluate? *New Directions for Evaluation*, 2000(87): 35-45.

White, H. (2005) *Challenges in evaluating development effectiveness*. IDS Working Paper 242, Brighton: Institute of Development Studies (IDS).

White, H. (2010) A Contribution to Current Debates in Impact Evaluation. *Evaluation*, 16(2): 153-164.

White, H. and Phillips, D. (2012) *Addressing attribution of cause and effect in small n impact evaluations: towards an integrated framework*. 3ie Working Paper 15, New Delhi: International Initiative for Impact Evaluation (3ie).

White, H. (2013) *An introduction to the use of randomised control trials to evaluate development interventions*. 3ie Working Paper 9. New Delhi: International Initiative for Impact Evaluation (3ie).

White, H. and Raitzer, D.A. (2017) *Impact Evaluation of Development Interventions: A Practical Guide*. Manila: Asian Development Bank.

Williams, B. and Imam, I. (2007) *Systems Concepts in Evaluation: An Expert Anthology*. Point Reyes, CA: EdgePress.

World Bank (2005) *The logframe handbook: a logical framework approach to project cycle management*. Washington DC: World Bank.

World Bank (2016) *The Use of Political Economy Analysis in the Design and Implementation of Development Policy Financing*. Concept Note for IEG Category I Learning Product, Washington DC: Independent Evaluation Group, World Bank.

Yanguas, P. and Hulme, D. (2014) *Can aid bureaucracies think politically? The administrative challenges of political economy analysis (PEA) in DFID and the World Bank*. ESID Working paper 33, Manchester: Effective States and Inclusive Development Research Centre (ESID), University of Manchester.

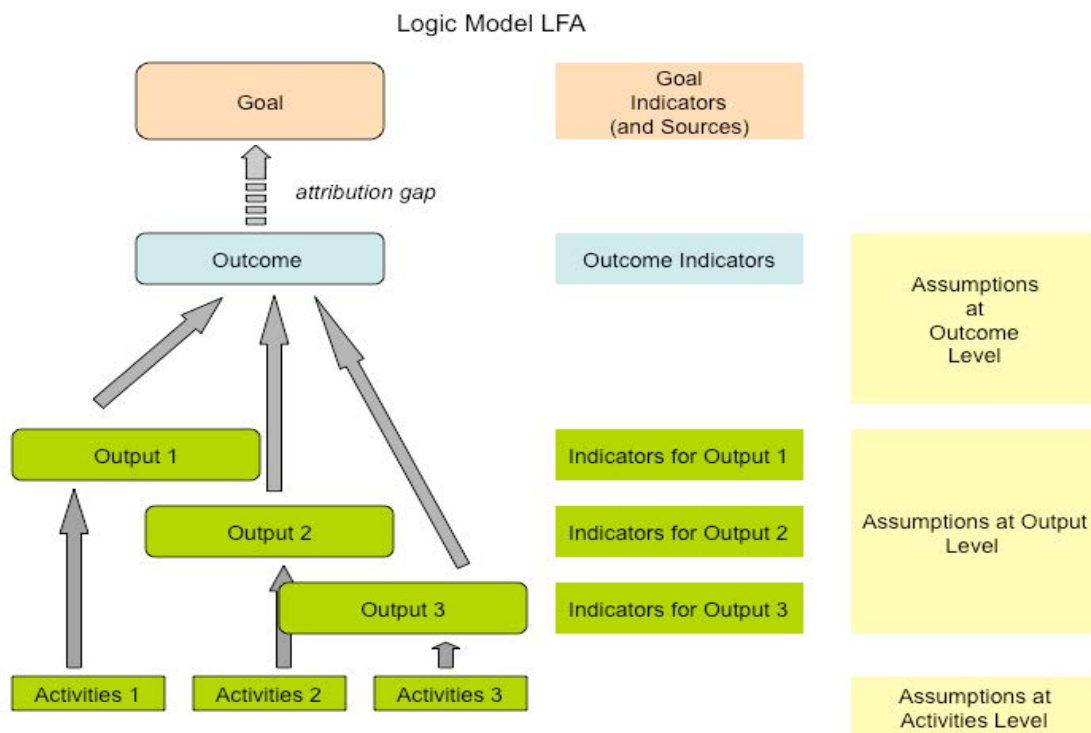
## Appendices

### Appendix 1: Glossary of Outcome Mapping terms (definition of elements and steps)

Outcome Mapping (tool for planning, monitoring and evaluation)		
<i>4 key planning questions: Why? Who? What? and How?</i>		
<b>Intentional design</b>	Vision	macro-level changes: improved human, economic, social and environmental wellbeing
	Mission	the focus of the program in support of the vision
	Boundary partners	Key actors, including individuals, groups and organizations whom the programme/project will target and interact with directly, and with whom the programme can anticipate contribution to influence and change
	Outcome challenges	Expected behavioural changes of a specific group of boundary partners and their relations to others if the programme/project achieves its full potential to facilitate changes
	Progress markers	a further breakdown of outcome challenge; A set of graduated indicators showing performance and progress of boundary partners towards the desired outcomes. They advance according to three different levels: 1) what one would expect to see as an early reaction/response to the programme/project's initial activities; 2) what one would like to see at a more advanced stage; and 3) what one would love to see as a transformative state of change if the programme/project were to have a significant influence
	Strategy maps	identify the strategies used by the programme to contribute to the achievement of an outcome; three types – causal, persuasive and supportive
	Organizational practices	<ol style="list-style-type: none"> <li>1. prospect for new ideas, resources and opportunities</li> <li>2. seek feedback from key informants</li> <li>3. obtain the support of next highest power</li> <li>4. assess and (re)design products, services, systems and procedures</li> <li>5. check up on those already served to add value</li> <li>6. share your best wisdom with the world</li> <li>7. experiment to remain innovative</li> <li>8. engage in organizational reflection</li> </ol>
<b>Monitoring</b>	Monitoring priorities	set the monitoring priorities on which data collection sheets are developed to track outcomes, strategies and or organizational practices
	Outcome journals	used for collecting data on the boundary partners' achievement of progress markers
	Strategy journal	used for collecting data on the programme's actions in support of the boundary partner
	Performance journal	used for collecting data on the organizational practices being adopted by the programme to remain relevant, sustainable and connected to its environment
<b>Evaluation planning</b>	Evaluation plan	selects and prioritizes a strategy, issue, or relationship to study and assess in depth; provides a short description of the main evaluation elements, and outlines the evaluation issue, the way findings will be used, the questions, the information sources, the evaluation methods, the evaluation team, the dates for the evaluation, and the approximate cost.

(Source: Earl et al., 2001)

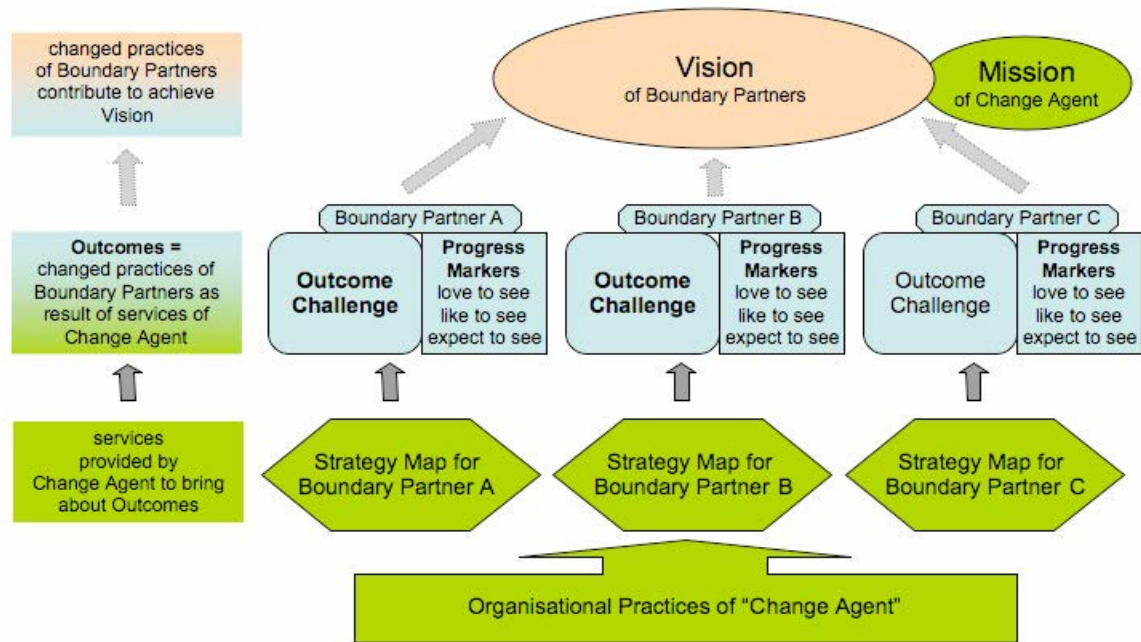
## Appendix 2: Logic model of the Logical Framework Approach



(Source: Roduner et al., 2008)

### Appendix 3: Logic model of Outcome Mapping

The Logic Model of Outcome Mapping



(Source: Roduner et al., 2008)

### Appendix 4: Framework used to collect progress-marker-related information

Progress marker information sheet							
Key actor (Boundary partner)					Location		
Date					Interviewer /Observer		
Outcome Challenge							
Progress Marker	Progress	Evidence	What happened? Description of what is said/observed	Influencing actors/factors	What next?	How about the past? (building baseline data retrospectively)	Other comment
Expect to see <u>Boundary partner A</u>							
1							
2							
3							
4							
Like to see <u>Boundary partner A</u>							
5							
6							
7							
8							
9							
10							
11							
12							
Love to see <u>Boundary partner A</u>							
13							
14							
15							

**Appendix 5: Interview guide – boundary partners: local consumer associations  
(Case 1)**

<b>Interview guide – boundary partners: local consumer associations (Case 1)</b>						
<b>General questions</b>						
1. Has the project 'Implementing Sustainable Consumption in Civil Society of Urban China' been successful from your perspective? If so (not), how and why?						
2. What has your organization gained from the project?						
3. Have there been any changes in the attitudes, practices, relationships or activities of your organization because of your participation in the project?						
4. If so, who or what changed – where and when?						
5. How has the project contributed to these changes, and what other actors and factors also influenced them?						
6. What challenges have you faced whilst participating in this project and what do you think can be improved in future?						
7. Did you know about the concept of sustainable consumption before the project began and did you undertake any related activities?						
8. What was the status of your organization in terms of promoting sustainable consumption and related practices?						
9. What are the most significant changes in your organization since beginning to participate in this project?						
10. What are your project follow-up plans?						
<b>Progress markers information</b>						
<b>Progress Markers</b>	<b>Progress</b>	<b>Evidence</b>	<b>What happened? Description of what is said/observed</b>	<b>Influencing actors/factors</b>	<b>What next?</b>	<b>How about the past? (building baseline data retrospectively)</b>
<b>Expect to see</b> <a href="#">local consumer associations</a>						
1. Signing voluntary agreements with larger retailers/supermarkets and SME suppliers						

2. Working with evidence-based communication campaigns often involving product tests, and putting pressure on suppliers and producers of harmful or non-sustainable products to change their supply or production approaches						
<b>Like to see</b> <a href="#">local consumer associations</a>						
3. Disseminating information on product quality comparison tests to consumers						
4. Including sustainable consumption websites developed during this project into the websites of consumer associations of the two cities						
<b>Love to see</b> <a href="#">local consumer associations</a>						
5. Raising awareness and strengthening capacity in helping mainstream citizens' sustainable consumption patterns and behaviour						
6. Transferring knowledge on best practices in sustainable consumption between the EU and China						

## Appendix 6: Interview guide – boundary partners: retailers (Case 1)

Interview guide – boundary partners: retailers (Case 1)						
General questions						
1. Has the project 'Implementing Sustainable Consumption in Civil Society of Urban China' been successful from your perspective? If so (not), how and why?						
2. What has your organization gained from the project?						
3. Have there been any changes in the attitudes, practices, relationships or activities of your organization because of your participation in the project?						
4. If so, who or what changed – where and when?						
5. How has the project contributed to these changes, and what other actors and factors also influenced them?						
6. What challenges have you faced whilst participating in this project and what do you think can be improved in future?						
7. Did you know about the concept of sustainable consumption before the project began and did you undertake any related activities?						
8. What was the status of your organization in terms of promoting sustainable consumption and related practices?						
9. What are the most significant changes in your organization since beginning to participate in this project?						
10. What are your project follow-up plans?						
Progress markers information						
Progress Marker	Progress	Evidence	What happened? Description of what is said/observed	Influencing actors/factors	What next?	How about the past? (building baseline data retrospectively)
<b>Expect to see</b> <a href="#">retailers</a>						
1. Signing voluntary agreements with local consumer associations						
2. Signing green supply contracts with SME suppliers to help establish product sustainability criteria by introducing green labels and providing information on energy costs						
<b>Like to see</b> <a href="#">retailers</a>						



3. Giving priority to buying sustainable products from SME suppliers						
4. Increasing percentage of green products on shelves						
5. Making efforts to influence suppliers and promoting choice editing						
<b>Love to see</b> <a href="#">retailers</a>						
6. Transferring knowledge on sustainable product labels and disseminating tools and skills to grow market demand for sustainable products						

## Appendix 7: Interview guide – boundary partners: suppliers (Case 1)

Interview guide – boundary partners: suppliers (Case 1)						
General questions						
1. Has the project 'Implementing Sustainable Consumption in Civil Society of Urban China' been successful from your perspective? If so (not), how and why?						
2. What has your organization gained from the project?						
3. Have there been any changes in the attitudes, practices, relationships or activities of your organization because of your participation in the project?						
4. If so, who or what changed – where and when?						
5. How has the project contributed to these changes, and what other actors and factors also influenced them?						
6. What challenges have you faced whilst participating in this project and what do you think can be improved in future?						
7. Did you know about the concept of sustainable consumption before the project began and did you undertake any related activities?						
8. What was the status of your organization in terms of promoting sustainable consumption and related practices?						
9. What are the most significant changes in your organization since beginning to participate in this project?						
10. What are your project follow-up plans?						
Progress markers information						
Progress Markers	Progress	Evidence	What happened? Description of what is said/observed	Influencing actors/factors	What next?	How about the past? (building baseline data retrospectively)
<b>Expect to see</b> <a href="#">suppliers</a>						
1. Signing voluntary agreements with local consumer associations						
2. Signing green supply contracts with target retailers						
<b>Like to see</b> <a href="#">suppliers</a>						
3. Promoting the development of green supply and procurement						

4. Taking innovative technical and managerial measures to produce more sustainable products						
5. Increasing the transparency of product information						
<b>Love to see</b> <a href="#">suppliers</a>						
6. Taking efforts to make their production chains more environmentally friendly from design through to recycling						

## Appendix 8: Interview guide – boundary partners: local authorities (Case 1)

Interview guide – boundary partners: local authorities (Case 1)						
General questions						
1. Has the project 'Implementing Sustainable Consumption in Civil Society of Urban China' been successful from your perspective? If so (not), how and why?						
2. What has your organization gained from the project?						
3. Have there been any changes in the attitudes, practices, relationships or activities of your organization because of your participation in the project?						
4. If so, who or what changed – where and when?						
5. How has the project contributed to these changes, and what other actors and factors also influenced them?						
6. What challenges have you faced whilst participating in this project and what do you think can be improved in future?						
7. Did you know about the concept of sustainable consumption before the project began and did you undertake any related activities?						
8. What was the status of your organization in terms of promoting sustainable consumption and related practices?						
9. What are the most significant changes in your organization since beginning to participate in this project?						
10. What are your project follow-up plans?						
Progress markers information						
Progress Markers	Progress	Evidence	What happened? Description of what is said/observed	Influencing actors/ factors	What next?	How about the past? (building baseline data retrospectively)
<b>Expect to see</b> <a href="#">local authorities</a>						
1. Keeping policy dialogue with project partners						
2. Seeking synergy between the project and related local programmes in the two cities						
<b>Like to see</b> <a href="#">local authorities</a>						
3. Supporting awareness raising and capacity building in consumer associations etc.						

4. Raising citizen awareness of sustainable consumption through support to training schools and relevant courses						
<b>Love to see</b> <a href="#">local authorities</a>						
5. Transferring knowledge on best practices in sustainable consumption between the EU and China						

### Appendix 9: Interview guide – project team staff (Case 1)

Interview guide – project team staff (Case 1)
1. Has the project ‘Implementing Sustainable Consumption in Civil Society of Urban China’ been successful from your perspective? If so (not), how and why?
2. What do you think are the most significant changes the project has brought?
3. From an actor-centred perspective, please describe changes in the behaviour, relationships, activities, actions, policies or practices of local consumer associations, retailers, suppliers or local authorities. State specifically what the change was, and when and where it took place.
4. How has the project contributed to these changes, and what other actors and factors also influenced them?
5. What challenges have you faced whilst involved in this project and what do you think can be improved in future?
<b>After the project:</b>
6. What are your next steps following this project?

### Appendix 10: Interview guide – donor representatives (Case 1)

Interview guide – donor representatives (Case 1)
1. What are the ideal outcomes the EU side expected from the project ‘Implementing Sustainable Consumption in Civil Society of Urban China’?
2. Has the project ‘Implementing Sustainable Consumption in Civil Society of Urban China’ been successful from your perspective and has it realized the expectations of the EU? If so, how and why?
3. What do you think are the most significant changes the project has brought?
4. From an actor-centred perspective, what do you think of the implementation process of the project? What about the project team? Have they implemented the project effectively and delivered satisfactory results? What do you think should be improved by each project partner? (those who you want to mention)
5. What challenges or risks have you faced whilst involved in this project and what do you think can be improved in future?
6. What method does the EU use to evaluate such a project, and how has it attributed achievement to the project in this case?
<b>After the project:</b>
7. What are your next steps following this project?

**Appendix 11: Interview guide – boundary partners: partner organizations  
(Case 2)**

Interview guide – boundary partners: partner organizations (Case 2 )						
General questions						
1. Has PRIME been successful from your perspective? If so (not), how and why?						
2. What has your organization gained from the project?						
3. Have there been any changes in the attitudes, practices, relationships or activities of your organization because of your participation in the project?						
4. If so, who or what changed – where and when?						
5. How has the project contributed to these changes and what other actors and factors also influenced them?						
6. Has the capacity building model of this project been effective? If so, how?						
7. Specifically, has your capacity, approach or behaviour in addressing the ultra-poor been built or changed whilst participating in the project (such as changes in the targeting mechanism, staff to client relationship, or innovation in adapting your approach to the ultra-poor as clients)?						
8. What challenges have you faced whilst participating in this project and what do you think can be improved in future?						
9. What was the status of your organization in terms of dealing with or providing services to the ultra-poor before joining the project?						
10. What are your project follow-up plans? Are you going to use the skills and approach learnt from PRIME in other work?						
Progress markers information						
Progress Markers	Progress	Evidence/ What happened? Description of what is said/observed	Influencing actors/factors (difficulties/challenges)	What next?	How about the past (building baseline data retrospectively)	Other comment
<b>Expect to see</b> <a href="#">partner organization A</a>						
1. Selecting appropriate beneficiaries, developing profile of them, forming groups, introducing a flexible saving system, and holding issues-based discussions in group meetings.						
2. Determining IGA-based credit needs, providing flexible credit facilities, and following up on the use of loans and debt						

situations of members						
3. Assessing loan needs of IGA loanees during the season of vulnerability and providing emergency loans						
<b>Like to see</b> <a href="#">partner organization A</a>						
4. Identifying training needs for viable IGAs of selected members, providing practice-oriented IGA training (skills, market linkage), and establishing linkage between members and markets						
5. Assessing technical service needs of IGA practitioners and providing/arranging necessary vaccinations						
6. Preselecting suitable trainees for selected informal vocational training and following up on provision of training						
7. Undertaking door-to-door primary health care services, delivering selected medicines, providing referral to specialized and high-level government organization, NGO and private facilities						
8. Facilitating the development of a post-disaster rehabilitation project for affected PRIME members, providing safe water and other supplies as available from government sources, and conducting cash-for-work activities to repair/maintain/develop community infrastructure						
<b>Love to see</b> <a href="#">partner organization A</a>						
9. Receiving management training to establish a viable and RBM-driven branch office, and providing skills training to respective staff members to implement PRIME components						



## Appendix 12: Interview guide – project team staff (Case 2)

Interview guide – project team staff (Case 2)
Interview questions
1. Has PRIME been successful from your perspective? If so (not), how and why?
2. What do you think are the most significant changes the project has brought to the partner organizations?
3. In terms of an actor-centred approach, have there been any changes in the attitude, relationships, activities or behaviour of PRIME partner organizations?
4. If so, what were the changes, and when and where did they take place?
5. How did the project contribute to these changes?
6. What other actors and factors also influenced these changes?
7. Has the capacity building model of this project been effective in terms of helping partner organizations become institutionally capable and sustainable micro-finance service providers? If so, how?
8. More specifically, do you think the capacity, approach or behaviour of partner organizations in addressing the ultra-poor has been built or changed whilst participating in the project (such as changes in the targeting mechanism, or staff to client relationship)?
9. What challenges have you faced whilst participating in this project (in general and also in terms of challenges related to partner organizations) and what do you think can be improved in future?
10. Are there any project follow-up plans such as further support to partner organizations to sustain their achievements?

### Appendix 13: Interview guide – staff from the donor agency (Case 2)

Interview guide – DFID staff (Case 2)
Interview questions
1. What are the ideal outcomes DFID expected from the PRIME project?
2. Has PRIME been successful from your perspective and has it realized the expectations of DFID? If so, how and why?
3. What do you think are the most significant changes the project has brought?
4. In terms of an actor-centred approach, do you think the PRIME model of capacity building has been effective for partner organizations in particular?
5. Do you think the project team has implemented the project effectively and delivered satisfactory results? What do you think should be improved?
6. What challenges or risks have you faced whilst involved in this project and what do you think can be improved in future?
7. What method does DFID use to evaluate such a project, and how has it attributed achievement to the project in this case?
8. What are your next steps following this project and what are DFID's expectations if it is extended?